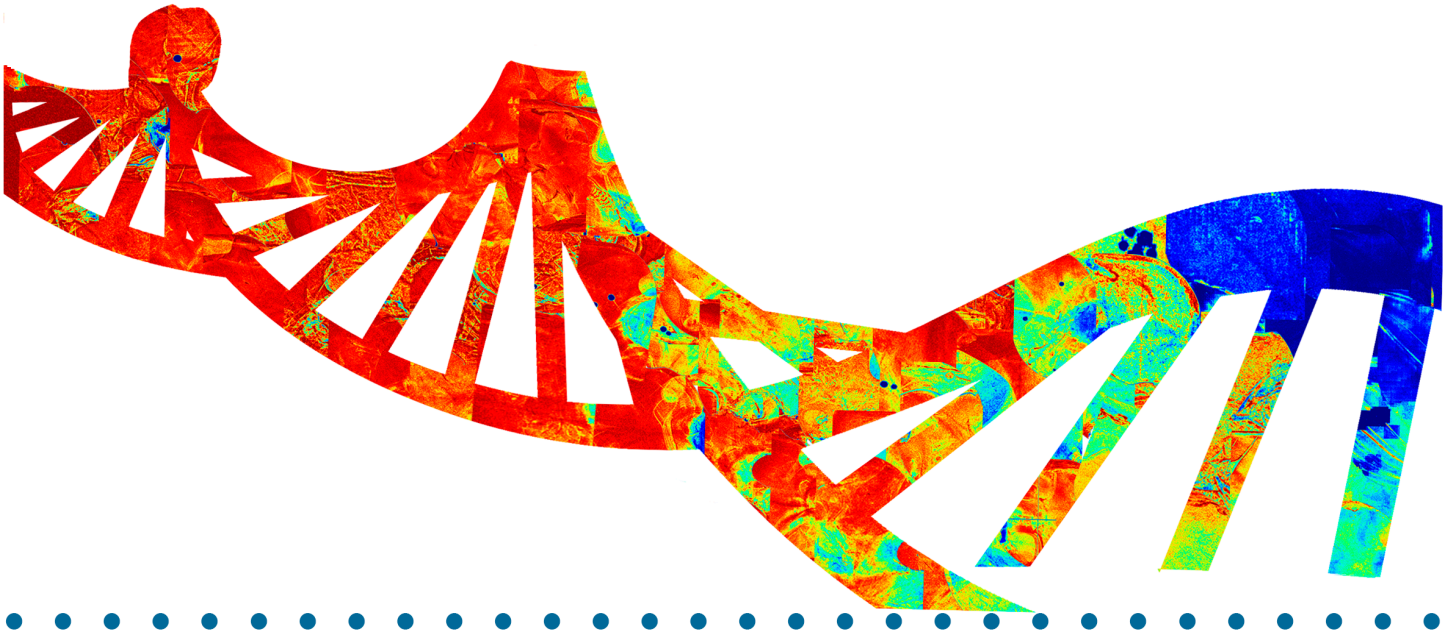
DECIPHERING INTRATUMOR MOLECULAR HETEROGENEITY IN GYNECOLOGICAL CANCER PROGRESSION

Alba Mota Jiménez

Madrid, 2017



TESIS DOCTORAL

DEPARTAMENTO DE BIOQUÍMICA
UNIVERSIDAD AUTÓNOMA DE MADRID

UNIVERSIDAD AUTÓNOMA DE MADRID
DEPARTAMENTO DE BIOQUÍMICA

DECIPHERING INTRATUMOR MOLECULAR HETEROGENEITY IN GYNECOLOGICAL CANCER PROGRESSION

ALBA MOTA JIMÉNEZ
TESIS DOCTORAL
MADRID, 2017

DEPARTAMENTO DE BIOQUÍMICA
FACULTAD DE MEDICINA
UNIVERSIDAD AUTÓNOMA DE MADRID

DECIPHERING INTRATUMOR MOLECULAR HETEROGENEITY IN GYNECOLOGICAL CANCER PROGRESSION

Memoria que presenta para optar al título de Doctor por la Universidad Autónoma de Madrid la Licenciada en Bioquímica

Alba Mota Jiménez

Directora de Tesis:

Dra. Gema Moreno Bueno

Profesora Titular de la Facultad de Medicina de la UAM

La presente tesis ha sido realizada en la Facultad de Medicina de la UAM y en el Instituto de Investigaciones Biomédicas "Alberto Sols" (CSIC-UAM) de Madrid, y financiada por una Beca de Formación de Profesorado Universitario (FPU) del Ministerio de Educación.

Dra. Gema Moreno Bueno

Profesora Titular de la Universidad Autónoma de Madrid

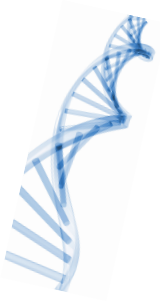
CERTIFICA que:

Alba Mota Jiménez, Licenciada en Bioquímica, ha realizado bajo su dirección en el Departamento de Bioquímica, Facultad de Medicina-Instituto de Investigaciones Biomédicas "Alberto Sols" (CSIC-UAM) de Madrid el trabajo titulado "**Deciphering intratumor molecular heterogeneity in gynecological cancer progression**".

El presente trabajo cumple, a nuestro juicio, con todos los requisitos necesarios para ser presentado y definido como Tesis Doctoral.

Madrid, 3 de marzo de 2017.

Fdo. Dra Gema Moreno Bueno



A mi familia,

A Iván

AGRADECIMIENTOS

"Duda siempre de ti mismo, hasta que los datos no dejen lugar a dudas"
Louis Pasteur

En primer lugar, agradecer a mi directora de tesis Gema Moreno la oportunidad que me dio al incorporarme en su grupo. Han sido unos años muy provechosos, de proyectos interesantes y múltiples y buenas colaboraciones. Gracias por abrirme las puertas a la investigación traslacional.

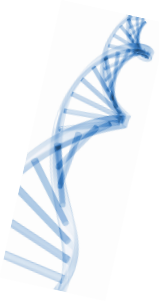
Gracias a mis compañeros de laboratorio, sería imposible resumir todo lo que tendría que agradeceros. En especial a Elena y María, juntas el Eje del Mal ha sido capaz de convertir en amistad lo que en su día llamábamos trabajo, nunca me faltará una sonrisa con vosotras. Al resto de mis compañeros en el MD Anderson. A Pablo por sus constantes correcciones y ocurrentes aportaciones (no siempre científicas, por supuesto). A Clara y a Eva, últimas pero esenciales incorporaciones. A las chicas de Fundación por su apoyo y sus desayunos, con pincho de tortilla a ser posible, sobre todo a Zaira por traducir mi manchego durante estos años. A las chicas de Biobanco y AP (en especial a Sandra) por su eterna sonrisa y disposición ante los miles de favores; así como a los patólogos del hospital, en especial a Alejandro y su paciencia infinita ante mis también infinitas peticiones. A los que con su paso temporal nos dejaron su huella, Julián, Heidys, y sobre todo Laura, esa sevillana salerosa que tanto me enseñó sobre cómo sobrellevar una tesis. A la gente del B16, que a pesar de la distancia física siempre han sido un gran apoyo, haciéndome sentir una más. A Marta y su incansable capacidad para escuchar, ¡creo que ya te tengo dicho todo!. A David, por su ayuda con las correcciones de última hora, el broche final a sus aportaciones a lo largo de todos estos años. Al sector femenino: María, Ángela, Lidia…mil gracias por vuestros favores administrativos, por vuestra ayuda continua y por vuestra alegría y apoyo. A todos los Cano, los actuales y los pasados, en especial a Celia y Vanesa, lo he pasado genial con vosotros. A mis compañeros en el laboratorio del MSKCC, que me acogieron como una más y generaron una red de amigos más que multicultural. Y a mis former roommates, fuisteis mi familia newyorkina.

A mis amigas de siempre, qué suerte haberos encontrado, ¡qué suerte vivir con vosotras en la *front row*! Gracias a Celia por su gran labor con la portada de esta tesis, gestada caña en mano, a ser posible un lunes. A los nuevos amigos que he ido conociendo a lo largo de mi vida académica. Los quimiquenses con los que empecé mi andadura universitaria estudiando electromagnetismo y buscando el gato de Schrödinger en cualquiera que sea el

bar donde quedemos. A mis compis del máster, mis amigos madrileños de corazón, el corredor de la tesis os espera, ¡no os escaparéis ninguno!

Muy especialmente a mi familia, que siempre apoyó mis decisiones y siempre confió en mi. A mis padres, a mi hermano y nueva hermana Cris, a mi abuela Antonia, mis tíos y primos, y a Adrián, la nueva incorporación que tantas alegrías nos ha dado en tan poco tiempo. También a los que me dejaron a lo largo del camino. A mi abuela Rosario, que siempre presumió de tener una nieta listísima (tanto que sabía hasta silbar). A mi abuelo Gabriel, que hasta el último día me preguntó si no me cansaba de estudiar y si de verdad lo que yo hacía era trabajar…aunque las dudas quedaban resueltas sabiendo que al menos me pagaban por ello. No puedo expresar la cantidad de cosas que he aprendido de todos vosotros, me habéis hecho ser quien soy.

Y por último, y por ello más importante, a Iván, mi compañero inseparable. Gracias por tu ayuda, tu confianza y tu inexplicable capacidad para hacerme reír sea cual sea el momento o el lugar. Te quiero.



RESUMEN/ABSTRACT

RESUMEN

Durante los últimos años, el desarrollo de técnicas de secuenciación masiva ha permitido la caracterización de la heterogeneidad intratumoral (HIT) en múltiples tipos de tumores, incluyendo los cánceres ginecológicos. En este sentido, aunque los tumores de ovario (COs) se han estudiado en profundidad, poco se sabe sobre los cánceres de endometrio (CEs). El CE es la neoplasia maligna ginecológica más común en los países desarrollados, sin embargo es el CO el que causa la mayor parte de las muertes asociadas a tumores ginecológicos. Esto es debido a que mientras que el CE se diagnostica principalmente en estadios tempranos, los COs suelen detectarse en estadios avanzados. Sin embargo, cuando el CE presenta metástasis al diagnóstico tiene un peor pronóstico, con una tasa de supervivencia global incluso menor que la de los COs.

Partiendo de estas premisas, los principales objetivos de la presente tesis implican la caracterización de la HIT en tumores ginecológicos, además de examinar la posible utilidad del análisis genético en biopsias alternativas como nueva herramienta diagnóstica que permita capturar la HIT en estos tipos de tumores, especialmente CE.

Como abordaje experimental se ha analizado la HIT mediante secuenciación masiva de exomas o dirigida en múltiples regiones del tumor primario y de la metástasis. Estos estudios se llevaron a cabo en un CO poco descrito hasta la fecha, caracterizado por mutaciones nulas en *TP53*, en el que también se identificaron las vías de señalización alteradas según su perfil mutacional, así como en varios CE metastáticos, incluyendo los subtipos endometrioide (CEE), seroso (CES) y ambiguo (CEA). Los resultados obtenidos en el análisis de CE revelaron diferentes grados de HIT a nivel genético, como así lo reflejan los múltiples patrones de evolución filogenética encontrados. Mientras que los CESs mostraron principalmente un patrón de monofilia en su evolución clonal, con las regiones metastáticas procedentes de un subclon ancestral, los CEEs presentaron patrones heterogéneos en su progresión filogenética. Además, la secuenciación de exomas de un caso de CEA nos permitió caracterizar molecularmente este subtipo de tumor así como los modelos animales (patient derived xenografts, PDX) derivados de él. Los resultados obtenidos fueron utilizados en un estudio *in silico* preliminar que permitió el testado de una terapia personalizada en dichos modelos.

Finalmente, mediante secuenciación masiva dirigida en aspirados uterinos (AUs), biopsias pre-operatorias mínimamente invasivas, observamos que su análisis representaba con precisión el perfil mutacional del tumor, incluso en muestras histológicamente no diagnosticables. Además, observamos que los AUs eran capaces de capturar la HIT encontrada en las regiones del tumor.

En conjunto, a lo largo de esta tesis doctoral hemos avanzado en la caracterización de la HIT en múltiples subtipos de cáncer ginecológico además de plantear nuevas alternativas diagnósticas basadas en el uso de la secuenciación dirigida como una solución parcial para superar el problema clínico derivado de la HIT, al menos en CEs.

ABSTRACT

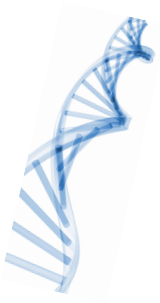
During the last years, the development of next generation sequencing technologies has shed light on intratumor heterogeneity (ITH) and clonal evolution in multiple tumor types. Regarding gynecological cancers, this has been widely described on ovarian tumors, although little is known about endometrial carcinomas (ECs). While EC is the most common gynecological malignancy in the developed world, tumors arising from the ovary account for the majority of deaths associated to gynecological cancers. The main reason of these differences is that whereas EC are mainly diagnosed at early stages, ovarian carcinomas are usually high grade tumors detected in advanced stages of the disease. However, ECs showing metastasis at diagnosis have a worse prognosis, with relative overall survival even lower than the observed in the ovary.

In this sense, the main objectives of the present thesis imply the characterization of ITH in gynecological cancers and the advance in the utility of the genetic analysis in alternative biopsies as new diagnostic tools to capture ITH in these types of tumors, especially EC.

In general, ITH was evaluated by primary tumor and metastasis multiregion studies based on targeted massive parallel or whole-exome sequencing (WES) analyses. These studies were performed in a less common ovarian carcinoma characterized by null mutations in *TP53*, where we also identified the potential signaling pathways altered according to the mutational profile, as well as in several metastatic ECs including endometrioid (EEC), serous (SEC) and ambiguous (AEC) histologies. Regarding EC, our results revealed different degrees of genetic ITH that were in turn reflected in the multiple phylogenetic evolution patterns found among the various histological and molecular subtypes. While SECs had mainly a monophyletic clonal progression, with metastatic regions arising from an ancestral subclone, EECs showed heterogeneous patterns of phylogenetic evolution. Additionally, WES allowed the molecular characterization of an AEC and the patient derived xenografts (PDXs) obtained from this case. These results were used for an *in silico* study that allowed the personalized treatment of these PDXs.

Finally, the use of targeted sequencing in uterine aspirates (UAs), low invasive pre-operative biopsies, revealed that UA accurately represents tumor mutational profile even in histologically non-diagnosable samples. Moreover, the genetic study of UA recapitulates the ITH found in the multiple tumor regions.

In conclusion, we have advanced in the characterization of ITH in multiple gynecological cancers subtypes and have introduced UAs to partially overcome the clinical problems derived from ITH, at least in ECs.



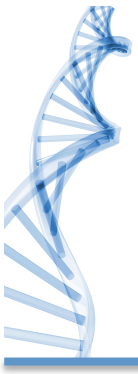
INDEX

INDEX OF TABLES

INDEX OF FIGURES

INDEX OF SUPPLEMENTARY MATERIALS: TABLES

INDEX OF SUPPLEMENTARY MATERIALS: FIGURES



ABBREVIATIONS

AH: Atypical Hyperplasia

AEC: Ambiguous Endometrial Carcinoma

CA-125: Cancer Antigen-125

CBPT: carboplatin

CCEC: Clear Cell Endometrial Carcinoma

CCF: Cancer Cell Fraction

CCOC: Clear Cell Ovarian Carcinoma

CGH: Comparative Genomic Hybridization

CGP: Cancer Genome Project

CS: Carcinosarcoma

CT: Computed Tomography

CTC: Circulating Tumor Cell

ctDNA: circulating tumor DNA

DNA: Deoxyribonucleic Acid

D&C: Dilation and Curettage

dNTP: deoxynucleotide mix

EC: Endometrial Carcinoma (en español, CE: carcinoma de Endometrio)

EEC: Endometrioid Endometrial Carcinoma

EOC: Endometrioid Ovarian Carcinoma

ER: Estrogen Receptor

FFPE: Formalin-Fixed Paraffin-Embedded

FIGO: International Federation of Gynecology and Obstetrics

FISH: Fluorescence in situ hybridization

H&E: Hematoxylin and Eosin

HGSOC: High Grade Serous Ovarian Carcinoma

HNPCC: Hereditary Non-Polyposis Colorectal Cancer

HQ: High Quality

HRR: Homologous Recombination Repair

ICGC: International Cancer Genome Consortium

IHC: Immunohistochemistry

ITH: Intratumor Heterogeneity (en español, HIT: Heterogeneidad Intratumoral)

Indel: Insertion or deletion of bases in the DNA

ISCIII: Instituto de Salud Carlos III

ISP: Ion Sphere Particle

LGSOC: Low Grade Serous Ovarian Carcinoma

LOH: Loss Of Heterozygosity

LQ: Low Quality

MAF: mutant allel frequency

MDR: Mutation Discovery Rate

MMMT: Malignant Mixed Mullerian Tumor

MMR: Missmatch Repair

MOC: Mucinous Ovarian Carcinoma

MRI: Magnetic Resonance Imaging

MSI: Microsatellite Instability

NAH: Non-Atypical Hyperplasia

NE: Normal Endometrium

NEEC: Non-Endometrioid Endometrial Carcinoma

NGS: Next-generation Sequencing

OC: Ovarian Cancer (en español, CO: cáncer de ovario)

OCT: Optimal Cutting Temperature

PBS: Phosphate-buffered saline

PCR: Polymerase Chain Reaction

PET: Positron Emission Tomography

PGM: Personal Genome Machine

PR: Progesterone Receptor

PTX: paclitaxel

SCNA: Somatic Copy Number Alteration

SEC: Serous Endometrial Carcinoma

SNP: Single Nucleotide Polymorphism

SNV: Single Nucleotide Variant

TCGA: The Cancer Genome Atlas

Tm: melting Temperature

TMA: Tissue Microarray

TR: Tumor Region

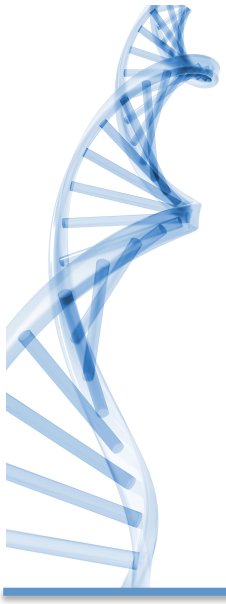
UA: Uterine Aspirate (en español, AU: Aspirado Uterino)

UTR: Untranslated Region

WES: Whole-Exome Sequencing

WGS: Whole-Genome Sequencing

WT: Wild-Type



INTRODUCTION

1 INTRODUCTION

1.1 Next-generation sequencing in cancer: unraveling intratumor genetic heterogeneity

1.1.1 Cancer genome sequencing: a bit of history

The necessity of unraveling the human genome sequence to understand the implication of genetic changes in cancer was proposed by Dulbecco as early as 1986[1]. However, it was not until 2001 that the International Human Genome Sequencing and Celera Consortiums published the first drafts of the human genome reference[2, 3]. Nowadays gene mutations and genome instability are considered one of the hallmarks of cancer[4]. During the period in which human genome was sequenced, different strategies based on cloning and first generation sequencing identified a large number of cancer related genes, including the most recurrent oncogenes and tumor suppressors[5]. The introduction of massively parallel sequencing around 2005[6, 7] revealed the feasibility of sequencing whole normal and tumoral genomes at a reasonable cost and timeframe[8, 9]. This involved the generation of an unprecedented volume of sequencing data and the necessity of overcoming huge bioinformatics challenges in terms of data storage, quality control, alignment, assembly and annotation. During the last years, the development of computational tools has allowed the implementation of next-generation sequencing (NGS) in routine research. While the first studies analyzed a short number of patients[10], big consortia as The Cancer Genome Atlas Research Network (TCGA), the International Cancer Genome Consortium (ICGC) and the Cancer Genome Project (CGP) among others have allowed the characterization of hundreds of tumors of different types[11, 12]. These studies have shown that few genes are frequently mutated in each cancer type, confirming those previously described as well as revealing new ones[13]. However, the number of low frequency mutated genes has substantially increased, suggesting the presence of genetic heterogeneity across tumors with the same origin (known as inter-tumor heterogeneity). Furthermore, recent studies are moving from whole-exome sequencing (WES) to whole-genome sequencing (WGS), since it has been shown that mutations in non-coding regions may have direct tumorigenic effects or lead to genetic instability[14, 15]. Additionally, a meta-analysis considering about 7,000 primary tumors of 30 different cases analyzed by WES and WGS, revealed that the combination of somatic mutations found in a tumor lead to a specific 'mutational signatures' and identified 21 recurrent cathegories[16]. While some of these signatures were related to the age of the patient, the exposure to mutagenics or defects in DNA repair, the

biological processes that originate the majority of them remain unknown. This study highlighted the diversity of mutational processes underlying cancer development and showed a potential utility in order to understand cancer aetiology.

1.1.2 Applications of NGS in cancer: use in the clinic

The widespread characterization of cancer genomes has increased the number of clinically relevant biomarkers for cancer-risk assessment, diagnosis and treatment, including the tailoring of therapeutic strategies based on actionable molecular alterations and resistance mechanisms[17]. However, large-scale genome sequencing studies are still unaffordable not only from an economical point of view, but also because of the limitations to apply their results into the clinic. It is important to note that few of described mutations have been functionally validated and the prediction of their consequence continues being a real challenge nowadays.

Nowadays, the implementation of NGS technologies into the clinic is mainly based on targeted sequencing of specific genes panels, due to the cost and the complexity of data analysis are sifnigicantly lower. A clear example of this was indicated in the 2015 National Comprehensive Cancer Network guidelines, which recommends the use of NGS gene panels for patients with hereditary ovarian cancer without mutations in the high-penetrance genes[18]. The use of this kind of panels allows the simultaneous analysis of multiple genes in several samples with a low DNA input and a high sensitivity. Additionally, these platforms can be applied in the analysis of formalin-fixed paraffin-embedded (FFPE) samples, favouring their utility in the clinical setting. Nevertheless, the selection of suitable genes for panel design, the need of additional validation, together with long-term storage and retrieval of data are still challenging[19].

1.1.3 New challenges for NGS: unraveling intratumor genetic heterogeneity

Intratumor phenotypic heterogeneity has been observed by pathologists since the early days of cancer knowledge, which led to propose the existence of a genetic heterogeneity implicated in the clonal evolution of tumors[20-22]. However, it has not been until the last years, with the development of NGS technologies, that intratumor heterogeneity (ITH) at a genomic point of view has been well demonstrated and deeply characterized. Initial sequencing studies comparing subsets of regionally[23, 24] and temporally[25] separated areas from the same tumor confirmed the existence of ITH not only in the primary lesions but also in metastatic regions. In fact, ITH has been described in numerous solid tumor types, including pancreatic[24, 26], renal[25, 27], lung[28-30], breast[31-33], colorectal[34, 35], glial[36-38] and prostate[39, 40] cancers, among others; as well as hematologic malignancies[41-46]. Several

studies have been also reported about gynecological cancers, including ovarian and endometrial carcinomas[47]. Since these types of cancer are the focus of this thesis, they will be discuss in depth in the next sections ***1.3.1.10*** and ***1.3.2.9***.

The analysis of multiple samples along tumor development has revealed different phylogentic evolution patterns. Thus, some tumors showed a monophyletic evolution[26, 48-50], with metastatic regions more closely related among them than to the primary tumor, but polyphyletic patterns were also found[24, 40] and revealed that distinct metastatic samples could arise from different subclones of the primary tumor.

The clinical implication of ITH remain a controversial issue nowadays, being necessary further studies to analyze its real impact on cancer progression, risk of relapse and treatment response and resistance[51]. Nevertheless, recent studies have suggested that ITH could be an independent prognostic factor of disease progression and survival[43, 52], although the underlying mechanisms are yet to be determined. Additionally, it is a fact that ITH also supposes a challenge for cancer diagnosis. Considering that molecular characterization of multiple tumor regions is not affordable in a diagnostic laboratory, supplementary tools should be developed to overcome this problem. In this sense, the use of alternative samples as liquid biopsies, such as circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs), has been proposed as a novel approach to capture ITH[53-55].

1.2 Female reproductive system

The female reproductive system is composed of internal and external sex organs[56]: i) internal organs are those included within the pelvis, counting the vagina, uterus, fallopian tubes and ovaries; and ii) external organs or genitals make up the vulva, which is connected to the uterus at the cervix (**Figure 1**).

Ovaries are the female gonads as well as endocrine glands, located alongside the lateral wall of the uterus. They are composed by an epithelial layer that covers the cortex (outer part consisting of ovarian follicles and stroma) and the medulla (inner part composed of supporting stroma and neurovascular structures).

The uterus is a hollow muscular organ in which the fetus develops during gestation. Anatomically, the uterus can be divided into fundus, corpus and cervix. The walls of the uterus are made up of three tissue layers:

- Perimetrium: external thin layer composed of a serous membrane continuous with the peritoneum that protects the uterus.
- Myometrium: middle layer formed by three smooth muscle fibers layers, responsible for uterine contractions.

- Endometrium: inner mucous epithelial membrane, highly vascularized, that lines the hollow lumen of the uterus. This layer changes throughout the menstrual cycle and it becomes thick and rich of blood vessels when prepare for pregnancy.

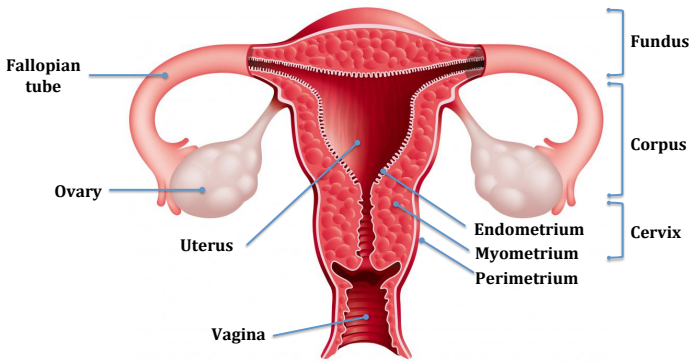


Figure 1 Female reproductive system. Scheme of the female reproductive system. The name of internal organs are indicated in the left side. Uterus parts are described in the right side. Figure adapted from http://www.wisegeekhealth.com.

1.3 Gynecological cancers

Gynecological cancers are those originated in the female reproductive organs, including uterine, ovarian, cervical, vulvar and vaginal cancers[57]. Although cervical cancer is the most frequent gynecologic cancer in the world, uterine and ovarian cancers are the most common in the developed countries[58]. Regarding the American Cancer Society estimation, there will be approximately 105,890 new cases diagnosed and about 30,890 deaths from gynecologic cancers in the United States during 2017 (Table 1).

Table 1 Estimated new cases and mortality from gynecologic cancers in the United States during 2017

Cancer Type	New cases*	Deaths*
Uterine	61,380	10,920
Ovarian	22,440	14,080
Cervical	12,820	4,210
Vulvar	6,020	1,240
Vaginal	4,810	1,150

*Data obtained from American Cancer Society

1.3.1 Ovarian cancer

1.3.1.1 Epidemiology of ovarian cancer

Ovarian cancer (OC) is the second most common gynecological cancer in developed countries[58]. About 3,300 new cases are diagnosed annually in Spain, representing 5% of female cancers behind breast, colorectal, uterine and lung cancer[59]. The 5-years relative

survival rate for these tumors is around 45%, although it strongly depends on the stage at diagnosis (**Figure 2**).

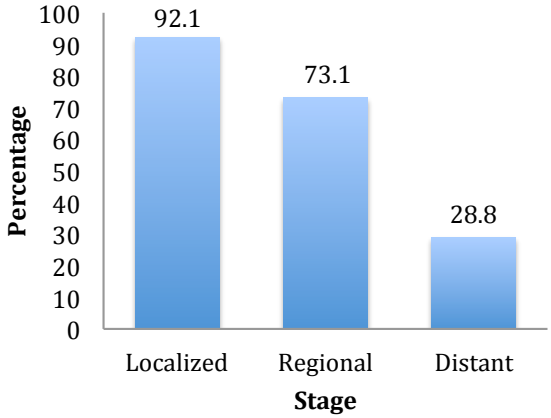


Figure 2 Five-year relative overall survival rate depending on stage at diagnosis in ovarian cancer. Percentage of cases with relative overall survival of at least five years depending on stage at diagnosis in ovarian cancer. Localized: tumor confined entirely to ovary. Regional: extended into surrounding organs or tissues. Distant: spread to parts of the body remote from the ovary. Adapted from http://seer.cancer.gov/.

While 92% of localized tumors survive 5 years before diagnosis, this percentage decreases to 29% in distant tumors[58]. Due to the majority of the patients are diagnosed at advanced stages, of which around 75% will recur after surgery and chemotherapy, OC is associated with a greater number of deaths, being the most lethal gynecological cancer.

1.3.1.2 Types of ovarian cancer

Ovarian cancers are divided into epithelial, germ cell and sex-cord stromal tumors, depending on the cell type that gives rise to the tumor. The majority of OCs (about 90%) are epithelial tumors, which start in the epithelial surface layer covering the ovary or in the distal fallopian tubes[60]. Germ cell and stromal tumors are rare subtypes, originated in the cells that generate the eggs or in the stromal cells that produce female hormones, respectively[57]. All these tumor subtypes can be subdivided into benign (non-cancerous), borderline (low malignant potential, without stromal invasion) or malignant tumors. Most of them are benign and confined into the ovary, while borderline and malignant tumors can spread to other parts of the body[57].

Epithelial ovarian tumors, also known as carcinomas, are the most common OCs. Based not only in histopathology characteristics but also in molecular features, five main types can be identified[61]: high-grade serous ovarian carcinomas (HGSOC), that are the most common accounting for about 70% of the cases, endometrioid (EOC, 10%), clear cell (CCOC, 10%), mucinous (MOC, 3%) and low-grade serous ovarian carcinomas (LGSOC, <5%). Their clinical and molecular characteristics are summarized in **Table 2**, and will be expounded on the next sections.

Table 2 Clinical and molecular features of the most common ovarian cancer types

	HGSOC	EOC	CCOC	MOC	LGSOC
Risk factors	*BRCA1/2* alterations	Lynch syndrome	Unknown	Unknown	Unknown
Precursor lesions	Tubal intraepithelial carcinoma	Atypical endometriosis	Atypical endometriosis	Cystadenoma, borderline tumor	Serous borderline tumors
Pattern of spread	Very early transcoelomic spread	Usually confined to pelvis	Usually confined to pelvis	Usually confined to ovary	Transcoelomic spread
Frequently mutated genes	*BRCA, TP53*	*PTEN, ARID1A*	*HNF1, ARID1A*	*KRAS, HER2*	*BRAF, KRAS*
Chemosensitivity	High	High	Low	Low	Intermediate
Prognosis	Poor	Favorable	Intermediate	Favorable	Intermediate

HGSOC: high-grade serous ovarian carcinoma, EOC: endometrioid ovarian carcinoma, CCC: clear cell ovarian carcinoma, MOC: mucinous ovarian carcinoma, LGSOC: low-grade serous ovarian carcinoma. Adapted from Prat *et a*l., 2012[61].

1.3.1.3 Risk factors in ovarian carcinoma

Several well-established risk factors have been identified in epithelial ovarian tumors. These include age, obesity, use of menopausal hormone therapy, endometriosis, diabetes and family history of breast and/or OCs, among others[62-67]. By contrast, the use of oral contraceptives, a late menarche, tubal ligation and hysterectomy, parity and breastfeeding have been associated with a lower risk of OC[68-72].

On the other hand, around 5 to 15% of the OC cases are related with hereditary syndromes (**Table 2**), the majority of them caused by germline mutations in *BRCA1* or *BRCA2* genes and related with HGSOC[73, 74]. Additionally, a smaller percentage of them, mostly with endometrioid histology, can be linked to other inherited conditions, such as Lynch syndrome, also known as hereditary nonpolyposis colorectal cancer (HNPCC)[75] (molecular basis are expounded in the *1.3.2.7* section).

1.3.1.4 Clinical manifestations and diagnosis

Ovarian cancer may cause several symptoms, including bloating, pelvic or abdominal pain, irregular menstruation or postmenopausal vaginal bleeding, urinary symptoms, loss of appetite and fatigue[59]. However, these symptoms appear more frequently when the disease has spread, which suppose a clinical problem.

The standard diagnosis protocol for OC begins with medical examination followed by imaging tests as ultrasounds, computed tomography (CT) scans, magnetic resonance imaging (MRI) and/or Positron emission tomography (PET) scans. A blood test can be also performed in order to detect cancer antigen-125 (CA-125), a protein used as a biomarker for ovarian tumors[76]. However, CA-125 has limited specificity since it can be increased in

other tumor types and even in benign conditions such as endometriosis, fibroids, pelvic inflammatory disease or pregnancy[77-79]. The final diagnosis will be always performed with a biopsy examination by the pathologist, using tissue or fluid in patients with ascites (fluid accumulated inside the abdomen)[80].

1.3.1.5 Ovarian cancer staging

Cancer staging classifies tumors regarding its extension, considering factors as primary tumor location, tumor size and cancer dissemination. There are numerous parameters for tumor staging. Although TNM classification of malignant tumor is the system most widely used for other types of cancer (T defines the size of the primary tumor and whether it has invaded nearby tissue, N indicates nearby lymph nodes that are involved and M refers to distant metastasis), FIGO (International Federation of Gynecology and Obstetrics) stage is recommended in gynecologic tumors typification. The last FIGO revision for ovarian, fallopian tubes and primary peritoneal cancers staging is summarized in Supplementary Table 1[57].

1.3.1.6 Standard treatment and new therapeutic approaches for ovarian cancer

The standard treatment in OC includes aggressive surgery followed by a combination of platinum and taxane chemotherapy[81]. However, although the initial response rate is higher than 80%, platinum resistance appears in approximately 25% of the patients within six months, and the majority of patients ultimately relapses due to chemo-resistant disease[82, 83]. The genetic and molecular factors that contribute to explain the high rate of resistance in this disease are currently unknown, although it has been shown that a defective status of the DNA repair machinery (especially *BRCA1/2* genes) correlates with a better response to platinum-based therapy[84].

In this sense, tumors with defects in DNA repair have been shown to be more sensitive to the inhibition of other DNA repair proteins as PARP[85]. This phenomenom, consisting of cell death caused by the simultaneous perturbation of two genes, is called synthetic lethality and its application supposed a real advance in OC treatment[86]. The OC research is currently focused on the extension of the use of different PARP inhibitors to other tumors with sensitizing molecular defects resulting in homologous recombination repair (HRR) deficiencies[87]. Furthermore, clinical trials including anti-angiogenic therapies to the standard treatment have shown a significant increase in the disease free survival time, suggesting the necessity of anti-angiogenic therapies maintenance until disease progression[88]. Nonetheless, effective predictors of response are required to select which

group of patients would benefit from these therapies, principally due to the substantial increase in the cost of treatment.

1.3.1.7 Molecular features of High Grade Serous Ovarian Carcinoma

As above pointed out, HGSOC represents the majority of OC cases. HGSOCs are widely characterized by *TP53* mutations, which appear in at least 96% of the cases[89, 90]. *TP53* is a key tumor suppressor gene implicated in G2 checkpoint, mediating cell cycle arrest, senescence or apoptosis in response to different kinds of cellular stress, including DNA damage, activated oncogenes or hypoxia[91, 92]. The majority of *TP53* mutations are missense substitutions in contrast to other tumor suppressor genes that present mainly truncating mutations[93]. This type of mutations leads to nuclear protein accumulation with a diffuse and strongly positive immunohistochemistry (IHC) staining. Nevertheless, approximately 30% of somatic *TP53* mutations are nonsense, frameshift or splicing junction variants that lead to complete absence of p53 protein, also known as null mutations. On the other hand, *TP53* wild-type tumors present a focal positive staining (less than 50% of the cells, usually less than 10%)[94] (**Figure 3**).

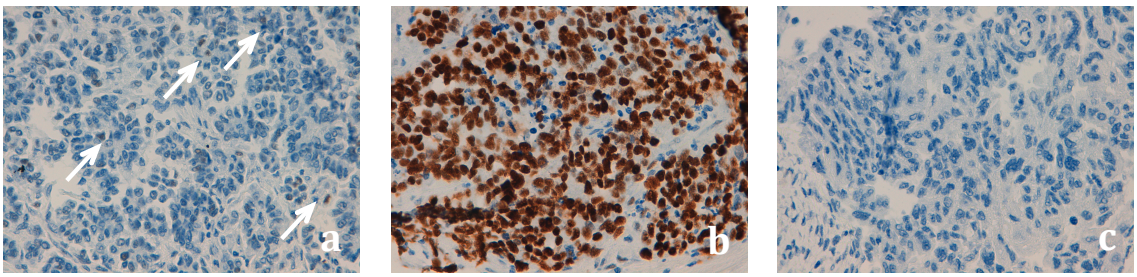


Figure 3 p53 immunostaining patterns in high-grade serous ovarian carcinomas. Immunostaining for p53 in high-grade serous ovarian carcinomas; representative examples for the different patterns of p53 staining. a) *TP53* wild-type: focal nuclear expression (marked with arrows) in less than 50% of tumor cells. b) *TP53* misense mutant: strong nuclear overexpression in more than 80% of tumor cells. c) *TP53* null mutant: complete absence of expression. Magnification 40x. Images were obtained from tumor tissue samples from MD Anderson Cancer Center, Madrid.

The majority of HGSOCs are also immunoreactive for *WT1* and p16 markers, as well as for Ki-67 which indicates a high proliferation index[95, 96]. Moreover, approximately two thirds of the cases express estrogen receptor (ER)[61].

As mentioned before, alterations in *BRCA1* and *BRCA2* genes, essential components of the HRR of DNA double-strand breaks, are also common in HGSOCs (**Table 2**). In fact, more than 15% of the cases present germline mutations, while somatic mutations or *BRCA1* promoter methylation occur in an additional 14-22% of the cases[90, 97]. Mutations in other

members of the HRR pathway, as *PALB2*, *RAD51*, *RAD50*, *BARD1*, *CHK2* and *BRIP1* have been also detected, although in a lower proportion of cases[98].

However, the hallmark of HGSOCs is the chromosomal instability and widespread somatic copy number alterations (SCNAs), probably as a consequence of the DNA repair disorders due to *TP53* or *BRCA1/2* mutations[99]. The first and most common SCNAs were identified in *MYC* and *ERBB2* oncogenes[100, 101].

1.3.1.8 TP53: a prognosis marker in ovarian cancer?

Although somatic mutations in the *TP53* gene are the most frequent genetic alterations in human cancers, showing HGSOCs the highest frequency among all the solid tumors with a 96%[89, 90, 93, 102], the prognosis value of *TP53* alteration is a controversial issue. In this sense, numerous studies have tried to correlate *TP53* mutational status with different clinical parameters such as overall survival or therapy response, but results remain conflicting[103]. This could be partially explained taking into account that the vast majority of studies has performed IHC to asses p53 alterations, a technique prone to misclassify an important number of cases due to the difficulty in distinguishing wild-type and null tumors (**Figure 3**). Likewise, other analyses limited *TP53* sequencing to those exons that encode the DNA binding domain, or did not differentiate between biological consequences of the mutations[104]. However, recent studies support that tumors with *TP53* null mutations present a worse outcome compared to those in which *TP53* harbors mutations involving overexpression, not only in OCs but also in leukemia, breast, colorectal and head and neck cancers[94, 105-107]. Conversely, a recent analysis using TCGA data sustains that *TP53* wild-type tumors could have a worse prognosis than mutant tumors, though it is important to note that no differentiation between missense and null mutations was considered[108].

1.3.1.9 Omics and ovarian cancer: The Cancer Genome Atlas

The TCGA Research Network published in 2011 an integrated genomic analysis of HGSOCs including 489 untreated stage II-IV tumors and their corresponding normal DNA[90]. Analyses of mRNA expression, microRNA expression, promoter methylation and DNA copy number alterations were performed in the totality of the patients, while WES was carried out in 316 of them.

As expected, WES analysis showed that *TP53* is the most frequently mutated gene in HGSOCs (96% of the samples), whereas lower prevalence but statistically recurrent somatic mutations were found in eight further genes including *BRCA1* and *BRCA2* (22% of the cases, including germline and somatic mutations), *RB1*, *NF1*, *FAT3*, *CSMD3*, *GABRA6* and *CDK12* (2-6%)[90]. Moreover, 113 significant focal somatic copy number alterations (SCNAs) were also found, supporting the relevance of chromosomal instability in this type

of tumor. Focal amplifications in *CCNE1*, *MYC* and *MECOM* (detected in more than 20% of tumors) and deletions of *PTEN*, *RB1* and *NF1* (observed at least in 2% of the tumors) were identified. Furthermore, promoter methylation analysis showed 168 genes silenced by epigenetic events in HGSOCs comparing with normal controls, including *BRCA1* in more than 10% of the cases, as previously reported[97].

Transcriptional analysis was able to classify HGSOC into four tumor subgroups, including 'immunoreactive', 'diferenciated', 'proliferative' and 'mesenchymal' subtypes, although no significant differences in survival rate were found among them. Additionally, miRNA expression analysis identified three subtypes that partially overlapped with the mRNA results. In this case, one of the clusters presented a significantly longer survival time[90].

Interestingly, a system biology study identified five main altered signaling pathways in HGSOCs, comprising RB (67% of cases altered), PI3K/RAS (45% of cases altered), NOTCH (22% of cases altered), HRR (51% of cases altered) and FOXM1 signaling (84% cases altered).

Nonetheless, it is important to notice that samples analyzed by the TCGA study included untreated primary tumors[90]. Given that the majority of HGSOCs recur due to planitum resistance, a recent publication based on WES analysis compared primary refractory, resistant, sensitive and matched acquired resistant tumors to further investigate in this sense[109]. Inactivation of the tumor suppressor genes *RB1*, *NF1*, *RAD51B* and *PTEN* by gene breakage was shown to contribute chemotherapy resistance acquisition, while *CCNE1* amplification was associated with primary platinum resistance. Other events implicated in platinum resistance were germline *BRCA1/2* mutation reversion, loss of *BRCA1* promoter methylation, and overexpression of the drug efflux pump *MDR1*[109].

1.3.1.10 Intratumor heterogeneity in ovarian cancer

Intratumor heterogeneity (ITH) has been extensively described in OCs[110-118]. First studies based on loss of heterozygosity (LOH) data by microsatellite and single nucleotide polymorphism (SNP) analyses demonstrated widespread ITH in primary ovarian tumors, suggesting a monoclonal origin[110]. This process was also found between metastatic lesions, being clonally related with the primary tumor[111]. These studies proposed a model in which OCs have a common clonal origin, evolving to polyclonal tumors due to genetic divergence[110, 111]. The role derived from ITH in cisplatin resistance was also described by array comparative genomic hybridization analysis, showing pre-existing minor resistant clones even before treatment[112].

Further analyses using more sensitive techniques, including WES and targeted massive parallel sequencing, have led to a deeper understanding of ITH in OC with a single

nucleotide resolution. The majority of studies agreed on the presence of extensive genomic and transcriptomic ITH in OCs, showing different degrees of heterogeneity depending on the patient[113-117]. These analyses also subscribed the presence of subclones in the untreated primary tumor that would give rise to recurrent disease, although the possibility of metastasis-to-metastasis spread has been also proposed[115]. The major differences regarding samples from the same patients were found between distant metastases and ovarian tumors[118]. Interestingly, the quantification of ITH may have a predictive value, showing a decreased progression free survival and overall survival for patients with extensive ITH[115].

All of these findings confirm that ITH supposes a real challenge in OC OCs[110-118]. In this sense, the genetic analysis of ascites appear to be a way to overcome ITH, since most of somatic mutations, SCNAs and methylation patterns are represented in ascitic cells[116]. With respect to ctDNA, despite the fact that it has been proposed as a solution in other types of cancer[53, 119], preliminary analyses with this source of genetic material in OC allowed the detection of mutations but with high variations in mutation detection sensitivity among patients[113].

1.3.2 Uterine Cancer

1.3.2.1 Epidemiology of uterine cancer

Uterine cancer is the most common gynecological cancer in the developed world, with about 5,000 new cases per year in Spain[59, 120]. On the contrary to OCs, uterine cancer is usually diagnosed in early stages, so the global survival rate is higher for this type of tumors. In this sense, approximately more than 80% of the patients will survive after 5 years, although this percentage decreases to less than 20% in tumors in which extension to distant sites is detected at diagnosis (**Figure 4**).

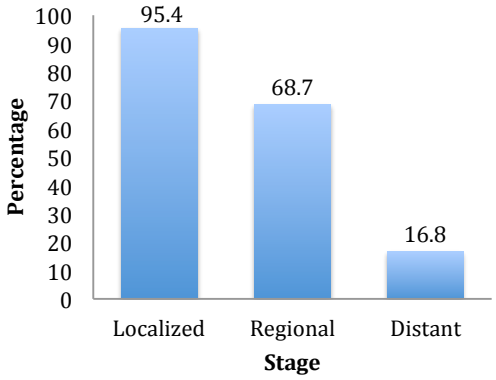


Figure 4 Five-year relative overall survival rate depending on stage at diagnosis in uterine cancer. Percentage of cases with a 5-year relative overall survival depending on stage at diagnosis in endometrial cancer. Localized: tumor confined to the uterus. Regional: extended into surrounding organs or tissues. Distant: spread to parts of the body remote from the uterus. Adapted from http://seer.cancer.gov/.

1.3.2.2 Types of uterine cancer

Uterine cancer is mainly divided into two subtypes regarding its cellular origin: uterine sarcomas and endometrial carcinomas[57]. Uterine sarcomas are originated in the muscle layer (myometrium) or supporting connective tissue of the uterus, representing less than 5% of uterine cancers. These mesenchymal tumors comprise leiomyomas, adenosarcomas, and endometrial stromal sarcomas[57]. The most common type of uterine cancer is endometrial carcinoma (EC), which originates in the epithelial cells of the endometrium. This thesis will be mostly focused on EC, given its high prevalence.

EC has been traditionally classified into two main groups with different clinical, pathological and molecular features[121, 122]. Type I or endometrioid endometrial carcinomas (EECs) are the most common ECs (about 75%), being mainly diagnosed in perimenopausal women. EECs are typically estrogen-related and low-grade tumors with good prognosis that coexist or are preceded by endometrial hyperplasia. In contrast, type II or non-endometrioid endometrial carcinomas (NEECs) are high-grade aggressive tumors associated with endometrial atrophy and poor prognosis, unrelated to estrogen and diagnosed in older women. These comprise several histological subtypes, being the most common serous (SEC) and clear cell (CCEC) carcinomas.

However, this dualistic model presents several limitations[123, 124]. For example, high-grade endometrioid carcinomas, accounting for 10-19% of type I EECs, seem to be more similar to type II cancers due their lack of association with endometrial hyperplasia and poor outcomes[125]. Moreover, there are tumors that present combined or mixed morphologic and molecular features. That is the case of carcinosarcoma (CS), also known as malignant müllerian mixed tumor (MMMT). CSs are uncommon biphasic neoplasms formed by malignant epithelial and sarcomatoid components (with the minor component representing at least 10% of the neoplasm), with high rates of recurrence and mortality[126, 127]. There are also some tumors with overlapping and intermediate characteristics between EECs and NEECs that do not show two distinct components, considered as ambiguous tumors (AEC)[128, 129].

Clinical and molecular characteristics of endometrial tumors are summarized in **Table 3**, and will be expounded on the following sections.

Table 3 Clinical and molecular features of the different endometrial carcinoma subtypes

	EEC		SEC	CCEC	CS
Clinical features					
Histological grade	Low	High	High	High	High
Metastasis	Uncommon	Lymph nodes Distant organs	Lymph nodes Peritoneal Distant organs	Lymph nodes Peritoneal -/+	Lymph nodes Peritoneal Distant organs
Prognosis	Favorable	Poor	Poor	Poor	Poor
Molecular characteristics					
DNA MMR loss	-/+	-/+	-	-/+	-
Common mutations	*PTEN, KRAS, PIK3CA, CTNNB1*	*PTEN, KRAS, PIK3CA, CTNNB1, TP53*	*TP53, PIK3CA*	*PIK3CA*	*TP53, PIK3CA*
ER/PR expresssion	+	-/+ (frequent)	-/+	-	-/+
p16 expression	-	-	+	-	-/+
Ki-67	Low	High	High	Low/High	High

EEC: endometrioid endometrial cancer; SEC: serous endometrial cancer; CCEC: clear cell endometrial cancer; CS: carcinosarcoma; ER: estrogen receptor; PR: progesterone receptor; MMR: mismatch repair. + : present/high; – : absent/low; –/+ : could be present or absent. Adapted from Murali *et al.*, 2014[123].

1.3.2.3 Risk factors in endometrial carcinoma

Endometrial cancer is associated with several risk factors, including age, overweight, obesity, and endometrial hyperplasia[130, 131]. Multiple risk factors associated to hormonal imbalance (due to high exposure to estrogens) have been described to increase uterine cancer risk, such as long menstrual lifespan, exposure to hormonal replacement therapy or nulliparity[132-134]. Medical disorders as diabetes, polycystic ovarian syndrome, and previous diagnosis of breast or ovarian tumors, which may involve tamoxifen treatment, are also linked to a higher risk[135-139]. Other factors as the use of oral contraceptives, multiparity and physical activity descrease the risk of suffering uterine cancer[140-142].

Approximately 2-5% of ECs are related to hereditary causes, most of them to Lynch syndrome, whose molecular basis are explained in the *1.3.2.7* section.

1.3.2.4 Clinical manifestations and diagnosis of uterine cancer

The most common clinical manifestation in uterine cancer is the abnormal vaginal bleeding or discharge, including bleeding after menopause or between periods. Pelvic pain, presence of a mass and loss of weight can also appear in later stages of the disease[143]. Diagnostic procedures in endometrial cancer are based on ultrasound test and especially endometrial tissue analysis. Pre-operative endometrial tissue can be obtained by dilation and curettage (D&C), hysteroscopy-guided biopsy or aspiration[144-146]. The D&C is nowadays considered a poor diagnostic procedure that may cause high discomfort in the

patients[144]. Conversely, uterine aspirate (UA) is a minimally invasive biopsy performed with a flexible, disposable suction curette or pipelle, preferentially a pipelle de Cornier. However, a failure rate of 8% has been reported in obtaining such samples, whereas 13% of the samples turn out to be histologically inadequate[146-148]. This especially occurs in postmenopausal women[149]. In fact, hysterecopy-guided biopsies have been reported to be more sensitive than those obtained by aspiration[145, 149, 150]. Nonetheless the use of hysteroscopy increases the risk of complications, such as uterine perforation, hemorrhage, potential damage to other organs, together with stroke and intoxication provoked by the dilation procedure[151]. Although there is no significant relation between hysteroscopy and the prognosis of the disease, it has also been associated with peritoneal disemination of tumor cells[152]. In spite of that, the diagnostic value of UAs was enlarged due its potential use in molecular diagnosis[153, 154].

1.3.2.5 Stages and prognosis factors in uterine cancer

Endometrial carcinoma staging is usually performed following the FIGO criteria as in OC (*1.3.1.5* section), although the TNM system is also employed. The main factors for an optimal staging include tumor location, depth of myometrial invasion and/or lymphatic invasion, and extra-uterine spread. Moreover, the most important prognosis factors in EC are not only defined by tumor staging, but tumor histology and grade are also decisive. The last FIGO revision for ECs is summarized in Supplementary Table 2[57].

1.3.2.6 Standard treatment and new therapeutic approaches for endometrial carcinoma

Nowadays the standard treatment of EC depends on tumor stage and histologic grade. This includes (alone, consecutively or in combination) surgery, radiation, hormone therapy and chemotherapy[155]. The use of adjuvant radio and/or chemotherapy is based on the presence of risk factors. However, the high rates of treatment fail in high aggressive and recurrent EC reveal the need of new therapeutic approaches. Novel targeted therapies have been proposed considering the molecular abnormalities described in EC (see *1.3.2.7* section)[156]. In this sense, current strategies are mainly focused in the inhibition of PI3K pathway through mTOR inhibitors, the use of EGFR and PARP inhibitors, and more recently, immunotherapies[157-161].

1.3.2.7 Molecular features of endometrial carcinomas

Molecular alterations in EC have been associated to the two types traditionally described, in which different gene expression and copy number profiles have been found[162-164]. Basically, EECs show microsatellite instability (MSI) and mutations in *PTEN*, *PIK3CA*, *KRAS*

and *CTNNB1*; while NEECs are characterized by p53 alterations, loss of heterozygosity (LOH) through extensive somatic copy number alterations (SCNAs) and less frequent alterations as *ERBB2* amplifications or *CDKN2A* (p16) mutations and overexpression[124]. However, no mutations in any of these genes have been exclusively detected in either tumor subtype. The most common alterations found in EC and its prevalence in type I versus type II sporadic tumors are summarized in **Table 4**.

Table 4 Common molecular alterations reported in endometrial carcinoma

	Function	Alteration	Type I prevalence	Type II prevalence
MSI[165-167]	DNA repair	Methylation, mutation	20-30%	0-2%
PTEN[168, 169]	Tumor suppressor	Mutation, deletion, methylation, expression	52-78%	0-11%
PIK3CA[170, 171]	Oncogene	Amplification	2-14%	46%
		Mutation	36-52%	24-42%
PIK3R1[172]	Oncogene	Mutation	21-43%	0-12%
KRAS[173]	Oncogene	Mutation	13-26%	0-10%
CTNNB1[174, 175]	Oncogene	Mutation	25-38%	0-5%
ARID1A[176, 177]	Tumor suppressor	Mutation	25-48%	11-26%
TP53[173]	Tumor suppressor	Mutation	5-10%	80-90%
PPP2R1A[178]	Tumor suppressor	Mutation	5-7%	15-43%
Her2/neu *(ERBB2)*[179]	Oncogene	Amplification, expression	0%	18-80%
CDKN2A[180, 181]	Tumor suppressor	Mutation, methylation, expression	10%	10-40%
CDH1[182]	Tumor suppressor	Loss of expression	50%	80%

MSI: microsatellite instability. Adapted from Salvesen *et al.*, 2012[183] and Murali *et al.*, 2014[123].

1.3.2.7.1 Molecular alterations associated to type I EC

As mentioned before, molecular alterations described in EEC include MSI and somatic alterations in the genes mentioned above.

The presence of MSI was initially described in patients with Lynch syndrome, which presented inherited germline mutations in MMR genes, including *MLH1*, *MSH2*, *MSH6* or *PMS2*. By contrast, in sporadic endometrial tumors, the main cause of MSI is *MLH1* promoter hypermethylation, although mutations in MMR genes are also present[165-167].

On the other hand, alteration of the PI3K/AKT/mTOR pathway by activating mutations in *PIK3CA* or *PIK3R1* genes and inactivation of the tumor suppressor gene *PTEN* are

frequently found in this tumor subtype[184]. In fact, *PIK3CA* mutations seem to be associated with invasion and adverse prognosis factors, such as blood vessel invasion[185]. *PTEN* mutations have been also detected in atypical and non-atypical hyperplasia, considered precursos lesions, which indicates that these are early events in the development of EEC[161]. The same observation was found in activating *KRAS* mutations, considered in turn a bad prognosis factor[186, 187]. Alterations in Wnt/β-catenin include *CTNNB1* mutations and lack of E-caherin expression, though the second is more common in type II tumors[174, 175]. Coexistence between MSI and *PTEN* and *KRAS* mutations has been described, while *CTNNB1* mutations appear to be independent of them[124]. Mutations in *ARID1A*, a recently described tumor suppressor gene implicated in chromatin remodeling, have been also found in both EC types, although seem to be predominant in endometrioid and clear cell carcinoma histologies[176, 177]. Mutations in *ARID1A* frequently co-occur with alterations in the PI3K pathway, and are associated with its activation[188]. The correlation between *ARID1A* and tumor stages and its presence in 28% of metastatic ECs suggest its contribution to the progression of EECs[189].

1.3.2.7.2 Molecular alterations associated to type II EC

The main genetic and genomic alterations in NEEC are *TP53* mutations and chromosomal instability, characterized by widespread chromosomal gain and losses[90]. Less frequent mutations have been described in *PPP2R1A* and *CDKN2A* genes, implicated in the negative control of cell growth and division[178]. In addition, amplification and overexpression of *ERBB2*, a member of the epidermal growth factor (EGF) tyrosine kinase receptor family involved in cell proliferation, have been observed in serous and clear cell carcinomas[179].

1.3.2.8 Omics and endometrial cancer: The Cancer Genome Atlas

The description of classical molecular alterations found in EC was mainly based on targeted and unique gene or pathways studies. However, the development of high-throughput techniques has allowed the expansion of these studies to a large-scale molecular way[47]. Thus, WES analyses previous to the TCGA revealed mutations in new candidate genes, although their biological and clinical relevance have not been elucidated. A WES study including 11 endometrioid and two mixed carcinomas showed mutations in *ARID1A*, *IGFBP1*, *WNT11*, *HER3* and *RPS6KC1* genes[188]. On the other hand, three independent studies revealed frequent alterations in chromatin-remodeling and/or ubiquitin ligase complex genes in serous tumors, including *CHD4* and *FBXW7* mutations and *CCNE1* amplifications[190-192]. In this sense, our group have recently reported frequently mutations in chromatin remodelling and DNA repair genes[193]. Additional WES studies in less frequent uterine cancer types including sarcomas and undifferentiated tumors have

been recently carried out, revealing a low rate of point mutations in sarcomas and a high molecular heterogeneity in undifferentiated tumors[194, 195].

Nevertheless, the integrated analysis performed by the TCGA is the most complete and influential study to the date[196]. A total of 373 endometrial carcinomas, including endometrioid (n=307), serous (n=53) and mixed (n=13) carcinomas, were characterized at a genomic, transcriptomic and proteomic level. These analyses allow the stratification of EC into four distinct molecular subgropus: POLE/ultramutated, MSI/hypermutated, copy-number low/endometrioid and copy-number high/serous-like. The main characteristics of each subgroup are highlighted in **Figure 5**.

	POLE **(ultramutated, n=17)**	**MSI (hypermutated, n=65)**	**Copy-number low (endometrioid, n=90)**	**Copy-number high (serous-like, n=60)**
Hystology and grade	Endometrioid (grade 1-3)	Endometrioid (grade 1-3)	Endometrioid (grade 1-2)	Serous, mixed and high-grade endometrioid
SCNAs	Few SNCAs (clusters 1 and 2)	Few SNCAs (clusters1-3)	Few SNCAs (clusters1-3)	**Extensive SNCAs (cluster 4)** Frequent *MYC*, *ERBB2* and *CCNE1* amplifications, and alterations in *LRP1B*, *FGFR3/1* and *SOX17*
Mutation rate	232 x 10^{-6} mutations per Mb	18 x 10^{-6} mutations per Mb	2.9 x 10^{-6} mutations per Mb	2.3 x 10^{-6} mutations per Mb
Frequent mutations	***POLE***, *PTEN, PIK3R1, PIK3CA, FBXW7, KRAS*	***RPL22***, *KRAS, PTEN, PIK3CA, PIK3R1, ARID1A*	***CTNNB1***, *PTEN, PIK3CA, PIK3R1, ARID1A,*	***TP53***, *PIK3CA, PPP2R1A, FBXW7*
Other molecular characteristics	Increased frequency of C➔A transversions	*MLH1* promoter methylation		Infrequent *PTEN* and *KRAS* mutations
Prognosis	Favorable	Intermediate	Intermediate	Poor
mRNA expression	Immunoreactive	Hormonal (decreased *MLH1*, increased *PGR*)		Mitotic (increased *CCNE1*, *PIK3CA*, *MYC*, *CDKN2A*)
Protein expression	Cluster 1: ↑ *ASNS* and *CCNB1*	Cluster 2: ↑ p-Akt and ↓*PTEN*	Cluster 3:↑ *RAD50*	Cluster 4: ↑ p53
Methylation		MC1: heavily methylated (*MLH1* promoter methylation)		MC3: minimal DNA methylation

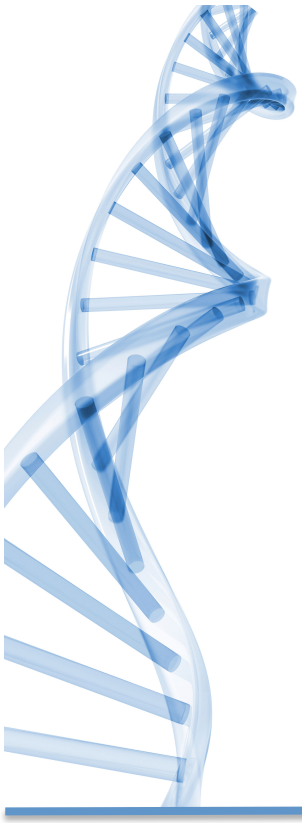
Figure 5 Molecular classification of endometrial carcinoma based on the TCGA study. Summary of the new molecular subgroups identified in the TCGA study[196]. The existence of four distinct molecular subgroups was proposed based on somatic copy number alterations (SCNAs), mutational profile and microsatellite instability (MSI). Additional classifications regarding mRNA and protein expression as well as DNA methylation were correlated with the integrated subgroups.

Additionally, novel classifications were also performed based on mRNA expression (n=333), protein expression (n=293) and DNA methylation (n=373). Interestingly, clusters obtained from mRNA sequencing, named 'mitotic', 'immunoreactive' and 'hormonal', strongly correlated with the integrated clusters explained above. Similar results were also found in the protein expression and DNA methylation analyses (**Figure 5**). Furthermore, these data were integrated to identify recurrently altered pathways, being RTK/RAS/β-catenin and PI3K pathways consistently affected. Finally, the TCGA study revealed a high genomic alterations concordance between high-grade serous ovarian carcinomas, serous endometrial carcinomas and basal-like breast carcinomas.

Moreover, a recent study has expanded and reanalyzed the TCGA data with improved bioinformatics methods[50], identifying a total of 49 frequently mutated genes, 21 of them not previously described in EC (Supplementary Table 3). Interestingly, some of them are involved in the estrogen receptor pathway (including *NRIP1* and *ESR1*).

1.3.2.9 Intratumor heterogeneity in endometrial cancer

Given the significance of ITH from the clinical point of view, it is important to study the relevance of this phenomenon in EC. Despite that the majority of large-scale sequencing analyses have been based on primary EC, a recent study including not only primary tumor but also pre-malignant and metastatic lesions have revealed the genomic evolution of EC[50]. In this analysis, atypical hyperplasia, considered as precursor tumoral lesions, showed at least one mutated gene, mainly *PTEN* or *PIK3CA*. However no recurrent metastasis-specific driver mutations were identified. Primary tumors and paired metastasis (n=26) showed similar levels of alterations, although only an average of 48% of mutations and 56% of SCNAs were common between paired samples. This ITH was also found in those cases in which multiple metastasis regions were analyzed (n=7). In the majority of these cases, the metastatic regions were more closely related to each other than to the primary tumor, suggesting a monophyletic evolution. Although bioinformatics analysis revealed the presence of subclonal populations in the primary tumor from which metastasis would arise, cases with multiple primary tumor regions were not included in this study to confirm this hypothesis. In this regard, additional studies covering multiple regions of the primary tumor are needed to better understand ITH in EC.



OBJECTIVES

2 OBJECTIVES

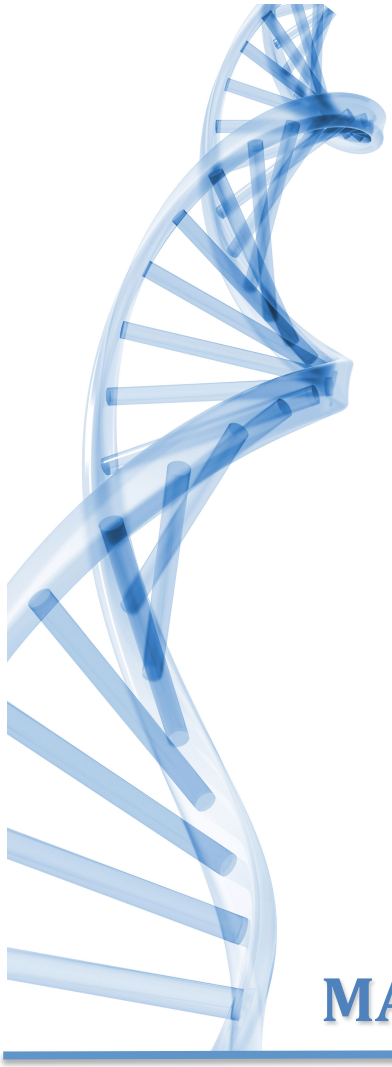
Previous studies have reported the presence of intratumor genomic and genetic heterogeneity (ITH) in gynecological cancers, mainly in ovarian carcinomas. However, little is known about endometrial carcinoma as well as the implications that ITH entails into the clinic. Due to the interest of our group in translational medicine and molecular diagnosis, as well as the application of new technologies based on next generation sequencing into the clinical practice, the following objectives were set for the present thesis:

1. Decoding ITH in the *TP53*-null high grade serous ovarian carcinoma subtype: multiple samples analysis in a patient.

2. Implementing new massive sequencing protocols for endometrial cancer diagnosis: mutational profile of uterine aspirates.

3. Characterizing ITH in endometrial carcinoma: phylogenetic evolution analysis regarding the mutational distribution in metastatic tumors.

4. Deciphering ITH of an ambiguous endometrial carcinoma by exome analysis: advancing in the translation of clinical research into the practice.



MATERIALS AND METHODS

3 MATERIALS AND METHODS

3.1 Human samples

For this thesis, different groups of formalin fixed paraffin embedded (FFPE) and frozen samples have been analyzed in three independent projects: WES in ovarian carcinoma, targeted massive sequencing in uterine aspirates and WES in metastatic endometrial carcinomas. Samples were collected at the MD Anderson Cancer Center (Madrid), the Arnau de Vilanova University Hospital (Lleida), the Vall d'Hebron University Hospital (Barcelona), the Bellvitge University Hospital (Barcelona), the Catalan Institute of Oncology (Barcelona) and the Medical University (Lublin). Tissue samples were obtained with the support of MD Anderson Foundation Biobank (record number B.0000745, ISCIII National Biobank Record), the Xarxa Catalana de Bancs de Tumors and Plataforma de Biobancos ISCIII (PT13/0010/0014, B.000609). Samples were retrospectively selected following the study criteria and the procedures established in the actual Spanish Law 14/2007 on Biomedical Research and the 1716/2011 Royal Decree on Biobanks. All the specimens were collected under patient informed consent and each project had the approval of Ethics and Scientific Committees of the corresponding institutions.

3.1.1 Ovarian samples

For the ovarian carcinoma study, a total of 12 samples were obtained from a 50-years-old patient diagnosed with a recurrent High Grade Serous Ovarian Carcinoma (HGSOC). The description of the samples is included in **Table 5**.

Table 5 Ovarian cancer samples characteristisc

Sample	Type of Tissue	Surgery	Anatomical location	Type of Study
P1	Frozen	Primary	Ovary	WES + Sanger
P2	FFPE	Primary	Rectum	Sanger
P3	FFPE	Primary	Right intra-pelvic dissemination	Sanger
P4	FFPE	Primary	Left intra-pelvic dissemination	Sanger
P5	FFPE	Primary	Thin mesentery	Sanger
P6	FFPE	Primary	Hepatic flexure	Sanger
IR1	Frozen	Recurrence	Pararectal dissemination	WES + Sanger
IR2	FFPE	Recurrence	Pararectal dissemination	Sanger
ER1	Frozen	Recurrence	Ileum	WES + Sanger
ER2	FFPE	Recurrence	Mesentery first jejunal loop	Sanger
ER3	FFPE	Recurrence	Ileum	Sanger
ER4	FFPE	Recurrence	Left sigmoid	Sanger
N	Frozen	Recurrence	Mesothelium	WES + Sanger

P: Primary Tumor, IR: intrapelvic recurrence, ER: extrapelvic recurrence, WES: Whole-exome sequencing, Sanger: Sanger sequencing

3.1.2 Endometrial samples

Two subset of patients were analyzed in two different projects of endometrial cancer (EC).

3.1.2.1 Uterine Aspirates project

In this study, the potential use of uterine aspirates (UAs) in genetic diagnosis as a low invasive biopsy was analyzed and compared with its paired hysterectomy sample. Cancer samples as well as pre-malignant and bening lesions were included (**Table 6**, **Table 7**).

Table 6 Endometrial samples analyzed in the uterine aspirate project

Diagnosis	n (%)
Endometrial Cancer	62/99 (61.4)
Atypical Hyperplasia	10/99 (9.9)
Non-Atypical Hyperplasia	7/99 (6.9)
Leiomyoma	7/99 (6.9)
Normal Endometrium	13/99 (12.9)

Table 7 Pathological features of the endometrial cancer samples analyzed in the uterine aspirate project

	n (%)
Tumor subtype (n=62)	
Endometrioid Endometrial Carcinomas	44 (71.0)
Serous Endometrial Carcinomas	9 (14.5)
MMMT	9 (14.5)
Endometrioid endometrial carcinomas: grade (n=44)	
1	12 (27.3)
2	8 (12.9)
3	24 (54.5)
Endometrioid endometrial carcinomas: FIGO stage (n=44)	
IA-B	32 (72.7)
IIA-B	6 (13.6)
IIIC	4 (9.1)
Serous endometrial carcinomas: grade (n=9)	
3	9 (100)
Serous endometrial carcinomas: FIGO stage (n=9)	
IA	2 (22.2)
IIA	1 (11.1)
IIIA-C	6 (66.7)
Carcinosarcomas: grade (n=9)	
3	9 (100)
Carcinosarcomas: FIGO stage (n=9)	
IA-B	6 (66.7)
II	2 (22.2)
IIIC	1 (11.1)
Carcinosarcomas: Epithelial component (n=9)	
Endometrioid	2 (22.2)
Serous	6 (66.7)
Clear Cell	1 (11.1)

3.1.2.2 WES in metastatic EC project

Multiple tumor regions from patients diagnosed with metastatic EC were analyzed in order to study intratumor heterogeneity (ITH). For this purpose, at least three samples from each tumor patient were analyzed by WES. A total of 11 EC cases (**Table 8**) were moleculary grouped according to the TCGA classification[196] as copy-number low (CN-low, n=3), microsatellite unstable (MSI, n=4) and serous like (n=4) carcinomas. Genetic variants detected in this analysis were validated by targeted massive parallel sequencing in the original samples as well as in additional samples from the same patients (**Table 9**).

Table 8 Summary of endometrial cancer cases analyzed in the WES project

Patient	Histology	Grade	FIGO	Myometrial infiltration	Lymph node invasion	POLE Status	MSI Status*	TCGA Classification**
EEC-1	Endometrioid	2	IIIC	100%	Yes (5/42)	WT	Negative	CN-low
EEC-2	Endometrioid	2	IIIC	75%	Yes (1/1)	WT	Negative	CN-low
EEC-3	Endometrioid	2	IB	80%	No	WT	Negative	CN-low
EEC-4	Endometrioid	2	IA	40%	No	WT	Positive-Low	CN-low
EEC-5	Endometrioid	3	IV	100%	-	WT	Positive-High	MSI
EEC-6	Endometrioid	3	IB	75%	No	WT	Positive-High	MSI
EEC-7	Endometrioid	3	IB	<50%	No	WT	Positive-High	MSI
SEC-1	Serous	3	IIIB	95%	-	WT	Negative	Serous-like
SEC-2	Serous	3	IIIA	80%	-	WT	Negative	Serous-like
SEC-3	Serous	3	IIIA	<50%	No	WT	Negative	Serous-like
AEC	Ambiguous	3	IIIC	95%	Yes (11/36)	WT	Negative	Serous-like

* MSI status was analyzed following the **3.8** section. **Molecular classification based on the previously published TCGA data[196], including MSI (hypermutated), copy-number low (CN-low, endometrioid) and copy-number high (serous-like) subgroups.

Table 9 Summary of primary tumor (T), metastatic (M) and normal (N) endometrial samples analyzed per patient in the WES project

Patient	Sample Name	Type of Tissue	Surgery	Anatomical location	Type of Analysis
EEC-1	**EEC-1_N**	Blood	Primary	Blood	WES+Validation
EEC-1	**EEC-1_T1**	Frozen	Primary	Fundus	WES+Validation
EEC-1	**EEC-1_T2**	Frozen	Primary	Left parametrium	WES+Validation
EEC-1	**EEC-1_M**	Frozen	Primary	Right pelvic node	WES+Validation
EEC-1	**EEC-1_T3**	FFPE	Primary	Left pelvic node	Validation
EEC-1	**EEC-1_T4**	FFPE	Primary	Uterus	Validation
EEC-1	**EEC-1_T5**	FFPE	Primary	Intern iliac node	Validation
EEC-1	**EEC-1_M2**	FFPE	Primary	Right primitive Node	Validation
EEC-2	**EEC-2_T1**	Frozen	Primary	Uterus	WES+Validation
EEC-2	**EEC-2_T2**	Frozen	Primary	Uterus	WES+Validation
EEC-2	**EEC-2_N**	Frozen	Primary	Non-tumoral tissue	WES+Validation
EEC-2	**EEC-2_M**	Frozen	Primary	Ovary	WES+Validation
EEC-2	**EEC-2_T3**	FFPE	Primary	Uterus	Validation
EEC-2	**EEC-2_T4**	FFPE	Primary	Uterus	Validation
EEC-2	**EEC-2_T5**	FFPE	Primary	Uterus	Validation
EEC-2	**EEC-2_T6**	FFPE	Primary	Uterus	Validation
EEC-2	**EEC-2_T7**	FFPE	Primary	Uterus	Validation
EEC-2	**EEC-2_M2**	FFPE	Primary	Ovary	Validation
EEC-2	**EEC-2_M3**	FFPE	Primary	Ovary	Validation
EEC-2	**EEC-2_M4**	FFPE	Primary	Ovary	Validation
EEC-3	**EEC-3_T1**	Frozen	Primary	Uterus	WES+Validation
EEC-3	**EEC-3_T2**	Frozen	Primary	Uterus	WES+Validation
EEC-3	**EEC-3_N**	Frozen	Primary	Non-tumoral tissue	WES+Validation
EEC-3	**EEC-3_M**	Frozen	Recurrence	Node	WES+Validation
EEC-3	**EEC-3_T3**	FFPE	Primary	Uterus	Validation
EEC-3	**EEC-3_T4**	FFPE	Primary	Uterus	Validation
EEC-3	**EEC-3_T5**	FFPE	Primary	Uterus	Validation
EEC-3	**EEC-3_M2**	FFPE	Recurrence	Iliac node	Validation
EEC-3	**EEC-3_M3**	FFPE	Recurrence	Iliac node	Validation
EEC-4	**EEC-4_T1**	Frozen	Primary	Uterus	WES+Validation
EEC-4	**EEC-4_T2**	Frozen	Primary	Uterus	WES+Validation
EEC-4	**EEC-4_N**	Frozen	Primary	Non-tumoral tissue	WES+Validation
EEC-4	**EEC-4_M**	Frozen	Recurrence	Epiplon	WES+Validation
EEC-4	**EEC-4_T3**	FFPE	Primary	Uterus	Validation
EEC-4	**EEC-4_M2**	FFPE	Recurrence	Peritoneum	Validation
EEC-4	**EEC-4_M3**	FFPE	Recurrence	Peritoneum	Validation
EEC-4	**EEC-4_M4**	FFPE	Recurrence	Peritoneum	Validation
EEC-5	**EEC-5_T1**	Frozen	Primary	Depth tumor	WES+Validation
EEC-5	**EEC-5_T2**	Frozen	Primary	Superficial tumor	WES+Validation
EEC-5	**EEC-5_N**	Frozen	Primary	Non-tumoral tissue	WES+Validation
EEC-5	**EEC-5_M**	Frozen	Primary	Peritoneun	WES+Validation
EEC-5	**EEC-5_T3**	FFPE	Primary	Uterus	Validation
EEC-5	**EEC-5_T4**	FFPE	Primary	Uterus	Validation
EEC-5	**EEC-5_T5**	FFPE	Primary	Uterus	Validation
EEC-5	**EEC-5_T6**	FFPE	Primary	Uterus	Validation
EEC-6	**EEC-6_T1**	Frozen	Primary	Depth tumor	WES+Validation
EEC-6	**EEC-6_T2**	Frozen	Primary	Superficial tumor	WES+Validation
EEC-6	**EEC-6_N**	Frozen	Primary	Non-tumoral tissue	WES+Validation
EEC-6	**EEC-6_M**	Frozen	Primary	Ovary	WES+Validation
EEC-6	**EEC-6_T3**	FFPE	Primary	Uterus	Validation
EEC-6	**EEC-6_T4**	FFPE	Primary	Uterus	Validation
EEC-6	**EEC-6_T5**	FFPE	Primary	Uterus	Validation
EEC-6	**EEC-6_T6**	FFPE	Primary	Uterus	Validation
EEC-6	**EEC-6_M2**	FFPE	Primary	Ovary	Validation
EEC-6	**EEC-6_M3**	FFPE	Primary	Ovary	Validation
EEC-6	**EEC-6_M4**	FFPE	Primary	Ovary	Validation
EEC-6	**EEC-6_M5**	FFPE	Primary	Ovary	Validation

Patient	Sample Name	Type of Tissue	Surgery	Anatomical location	Type of Analysis
EEC-7	**EEC-7_T1**	Frozen	Primary	Uterus	WES+Validation
EEC-7	**EEC-7_T2**	Frozen	Primary	Uterus	WES+Validation
EEC-7	**EEC-7_M**	Frozen	Recurrence	Diaphragm	WES+Validation
EEC-7	**EEC-7_N**	Frozen	Recurrence	Blood	WES+Validation
EEC-7	**EEC-7_T3**	FFPE	Primary	Uterus	Validation
EEC-7	**EEC-7_M2**	FFPE	Primary	Diaphragm	Validation
EEC-7	**EEC-7_M3**	FFPE	Primary	Diaphragm	Validation
EEC-7	**EEC-7_M4**	FFPE	Primary	Diaphragm	Validation
EEC-7	**EEC-7_M5**	FFPE	Primary	Diaphragm	Validation
SEC-1	**SEC-1_T1**	Frozen	Primary	Uterus	WES+Validation
SEC-1	**SEC-1_T2**	Frozen	Primary	Uterus	WES+Validation
SEC-1	**SEC-1_T3**	Frozen	Primary	Uterus	WES+Validation
SEC-1	**SEC-1_T4**	Frozen	Primary	Uterus	WES+Validation
SEC-1	**SEC-1_N**	Frozen	Primary	Non-tumoral uterus	WES+Validation
SEC-1	**SEC-1_M**	Frozen	Primary	Ovary	WES+Validation
SEC-1	**SEC-1_T5**	Frozen	Primary	Uterus	Validation
SEC-1	**SEC-1_T6**	Frozen	Primary	Uterus	Validation
SEC-1	**SEC-1_T7**	Frozen	Primary	Uterus	Validation
SEC-1	**SEC-1_M2**	FFPE	Primary	Ovary	Validation
SEC-1	**SEC-1_M3**	FFPE	Primary	Ovary	Validation
SEC-1	**SEC-1_M4**	FFPE	Primary	Tube	Validation
SEC-2	**SEC-2_T1**	Frozen	Primary	Uterus	WES+Validation
SEC-2	**SEC-2_T2**	Frozen	Primary	Uterus	WES+Validation
SEC-2	**SEC-2_T3**	Frozen	Primary	Uterus	WES+Validation
SEC-2	**SEC-2_N**	Frozen	Primary	Non-tumoral uterus	WES+Validation
SEC-2	**SEC-2_M1**	Frozen	Primary	Ovary	WES+Validation
SEC-2	**SEC-2_T4**	Frozen	Primary	Uterus	Validation
SEC-2	**SEC-2_M2**	Frozen	Primary	Ovary	Validation
SEC-2	**SEC-2_M3**	FFPE	Primary	Ovary and tube	Validation
SEC-2	**SEC-2_M4**	FFPE	Primary	Ovary	Validation
SEC-3	**SEC-3_T1**	Frozen	Primary	Uterus	WES+Validation
SEC-3	**SEC-3_T2**	Frozen	Primary	Uterus	WES+Validation
SEC-3	**SEC-3_M1**	Frozen	Primary	Ovary	WES+Validation
SEC-3	**SEC-3_M2**	Frozen	Primary	Ovary	WES+Validation
SEC-3	**SEC-3_N**	Frozen	Primary	Non-tumoral uterus	WES+Validation
SEC-3	**SEC-3_T3**	FFPE	Primary	Uterus	Validation
SEC-3	**SEC-3_T4**	FFPE	Primary	Uterus	Validation
SEC-3	**SEC-3_M3**	FFPE	Primary	Tube	Validation
AEC	**AEC_T1**	Frozen	Primary	Superficial tumor	WES+Validation
AEC	**AEC_T2**	Frozen	Primary	Superficial Tumor	WES+Validation
AEC	**AEC_N**	Frozen	Primary	Non-tumoral tissue	WES+Validation
AEC	**AEC_M**	Frozen	Primary	Node	WES+Validation
AEC	**AEC_T3**	FFPE	Primary	Uterus	Validation
AEC	**AEC_T4**	FFPE	Primary	Uterus	Validation
AEC	**AEC_T5**	FFPE	Primary	Uterus	Validation

N: normal, T: primary tumor, M: metastasis, WES: whole-exome sequencing, Validation: targeted massive parallel sequencing using Ion Torrent System

3.2 Samples processing

3.2.1 Tissue fixation and paraffin embedding

Tissue samples from surgery were immediately processed by the pathologist and fixed in neutral tamponated formalin (PanReac-AppliChem). Then, samples were included in paraffin blocks for their analysis and conservation, after several steps of dehydration with increasing ethanol (PanReac-AppliChem) concentrations. Paraffin blocks were cut with a RM2255 microtome (Leica, Germany) into sections of 4μm and counterstained with hematoxylin and eosin (H&E). Two independent pathologists evaluated the histology and tumor cell percentage of each sample.

3.2.2 Tissue freezing procedure

Tissue samples for frozen conservation were processed in less than 30 minutes since the surgical procedure. Representative zones of the tumor (about $5mm^2$) with viable tissue, avoiding ischemia areas, were selected by the pathologist. After resection, samples were included in a vinyl cryomold (Sakura, CA, USA), embedded in Optimal Cutting Temperature (OCT) (Sakura, CA, USA) and frozen by immersion in isopentane (VWR, PA, USA) (-56°C to -62°C). Frozen tissue blocks were stored at -80°C until use by the corresponding Biobank. Tissue sections (5-10μm) were obtained using a HM525NX cryostat (ThermoFisher, Germany) and counterstained with H&E to be evaluated by two independent pathologists.

3.2.3 Tissue Microarrays

Tissue microarrays (TMA) were performed with a 'Quick Ray' manual tissue microarrayer and premade recipient blocks with 120 holes (10x12) of 1mm diameter (IHC World, MD, USA). Two different representative tumor regions from each case were selected. The marked tissue was punched with a Quick-Ray needle (1mm) and delivered to the intended hole of the recipient block. After compleating all the samples, the recipient block was transfered to the embedding mold and incubated at 60°C for 30 minutes or until it turned completely transparent. Paraffin at 64°C was added to the block and the cassette array was placed on top of the paraffin. The block was solidified in a cold plate and cut (3-5μm) in a RM2255 microtome (Leica).

3.2.4 Uterine aspirate processing

Uterine aspirates were collected with a Cornier pipelle by clinicians and transported to the laboratory immediately. Samples were transferred to a 1.5ml Eppendorf tube and maintained on ice until processing (less than one hour). Aspirates larger than 700μl were

split up into two tubes after homogenization. Aspirate samples were diluted in a 1:1 ratio with PBS 1X and shook manually by inverting the tube several times. Then, they were centrifuged for 20 minutes at 2500g, 4°C. Supernatant and pellet were separated using a micropipette and both were stored at -80°C until use. In some of the cases, uterine pellets were processed as FFPE samples to generate blocks as previously described (see **3.2.1** section).

3.2.5 Immnohistochemisty analysis

Immunostaining was performed on 2-4µm FFPE tissue sections. After deparaffinization (60°C during 1 hour), PT Link module (Dako, Denmark) was used for antigen retrieval. Immunohistochemistry, amplification and visualization of immune complexes were performed in an Autostainer (Dako, Denmark), using EnVision FLEX+ (Dako, CA, USA). Primary antibodies used in the different projects are included in Supplementary Table 4.

3.3 DNA extraction and quantification

DNA was extracted from several FFPE or frozen tissue sections (10-15µm) and uterine aspirate pellets. Tissue was resuspended into Tris-HCl 50mM pH8 and digested with Proteinase K (10mg/ml of Proteinase K, recombinant, PCR grade, Roche), followed by phenol (Sigma-Aldrich) extraction and ethanol (PanReac-AppliChem) precipitation. Additionally, DNA from blood samples was obtained using DNeasy Blood & Tissue Kit (Qiagen), according to the manufacturer's protocols. DNA quantification was performed using NanoDrop 2000 Spectrophotometer and Qubit 3.0 Fluorometer with the Qubit dsDNA BR Assay Kit (ThermoFisher). DNA quality was evaluated using electrophoresis in agarose gel (Agarose D1 Low EEO, Conda-Pronadisa) and visualized with the Chemidoc XRS+ Molecular Image (BioRad).

3.4 Targeted massive parallel sequencing

Targeted massive parallel sequencing was performed using IonTorrent System (ThermoFisher) and following manufacturer's protocols. Sequencing workflow is summarized in **Figure 6**.

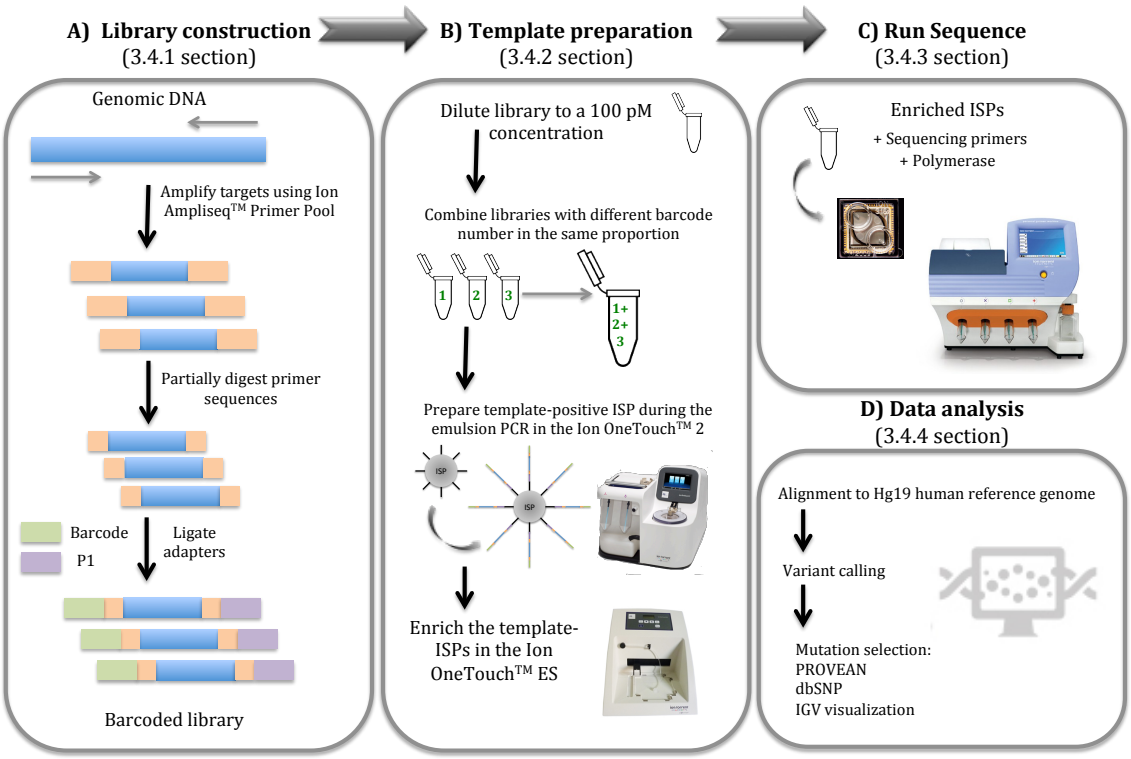


Figure 6 Scheme of targeted massive parallel sequencing protocol summary. Targeted massive parallel sequencing was performed using IonTorrent technology (ThermoFisher) and following manofacturer's protocol. Briefly, after library construction with Ion Ampliseq kit (A), a template with a mix of libraries were generated using Ion OneTouch 2 and Ion OneTouch ES (B). Samples were runned in the Ion PGM system and analyzed by the Torrent Suite Software v.4.2.1 (C).

3.4.1 Library construction

For library preparation, 10ng of DNA quantified according to Qubit protocol were used. Multiplex PCR for preparation of amplicon libraries was performed using the Ion AmpliSeq Library Kit 2.0 and Ion AmpliSeq Cancer Hotspot Panel v2 (https://www.thermofisher.com/order/catalog/product/4475346) or Custom Panels (Supplementary Table 5), which included the specific oligonucleotides designed for the genes of interest in each project. The following conditions were used in an AB 2720 Thermal Cycle (ThermoFisher): 99, 2min; cycle 99°, 2min/60°C, 4min; 10°C, hold. The number of cycles used in the PCR amplification depended on the tissue type and the number of primers used in the reaction (**Table 10**). After PCR amplification, primers were partially digested using FuPa reagent (50°, 10min; 55°, 10min; 60°, 10min; 10°, hold for up to 1 hour). Ion P1 Adapter and Ion Xpress Barcode X were combined and ligated to the amplicons (22°, 30min; 72°, 10min; 10°, hold), in order to identify each sample with a specific barcode. Libraries were purified and then quantified using the Ion Library TaqMan Quantitation Kit and ViiA 7 system (ThermoFisher). Fo this purpose, libraries were diluted (1:50) to reach a concentration within the range of the *E. coli* DH10B Control

Library standards. This Control Library is a validated, pre-quantified, ready-to-use control that was diluted serially (from 6.8pM to 0.00068pM), generating a standard curve to the qPCR. Relative concentration of the sample libraries to the Control Libraries were obtained from the qPCR analyses. Libraries were stored at -20°C in LoBind tubes (Eppendorf) until use.

Table 10 Number of cycles used in the multiplex PCR amplification for targeted massive parallel sequencing

Primer pairs per pool	Number of cycles	
	Standard DNA	FFPE DNA
12-24	21	24
25-48	20	23
49-96	19	22
97-192	18	21
193-384	17	20
385-768	16	19

3.4.2 Template preparation

Template preparation and enrichment were performed using Ion PGM Template OT2 200 Kit (uterine aspirates project) or Ion PGM Hi-Q OT2 Kit (EC WES project) and the Ion OneTouch 2 System. Libraries were diluted in nuclease-free water to a final concentration of 100pM and combined at the same proportion. Two µl of combined libraries were diluted in 23µl of water, obtaining a final volume of 25µl. Diluted libraries were used for the preparation of template-positive Ion PGM Hi-Q Ion Sphere Particles (ISPs) according to the manufacturer's protocol. In summary, an emulsion PCR was perfomed in the Ion OneTouch 2 followed by an enrichment executed by the Ion OneTouch ES. Enriched ISPs could be stored at 2°C to 8°C for up to 3 days.

3.4.3 Run Sequence

Ion PGM Sequencing 200 Kit v2 (uterine aspirates project) or Ion PGM Hi-Q Sequencing Kit (EC WES project) and Ion PGM System (ThermoFisher) were used for Next-Generation Sequencing according to the manufacturer's protocols. Enriched ISPs were centrifuged and mixed with sequencing primers and sequencing polymerase. A range of 12-16 libraries were run in 316 or 318 v2 chips to reach the optimum coverage, with at least 100 reads per base (100x).

3.4.4 Bioinformatics analysis

Alignment to Hg19 human reference genome and variant calling were performed using Torrent Suite Software v.4.2.1 (ThermoFisher). All samples were sequenced and analyzed in equal conditions. Variants with a Phred-score quality value less than 100 were

considered as low quality (LQ) variants (http://129.130.90.13/ion-docs/Technical-Note---Quality-Score_6128102.html). Prediction of the genomic variant impact on the biological function of the corresponding protein was performed using PROVEAN (Protein Variation Effect Analyzer) Genome Variants tool (http://provean.jcvi.org/index.php)[197]. Genetic variants with a damaging or deleterious consequence predicted by at least one of the PROVEAN predictors (PROVEAN[197] or SIFT[198]) were selected and visually checked with the Integrative Genomics Viewer (IGV) v.2.3.40, Broad Institute[199]. Variants with global minor allele frequencies higher than 0.05 were considered Single Nucleotide Polymorphisms (SNPs) and were rejected (data from dbSNP, http://www.ncbi.nlm.nih.gov/SNP/).

3.5 Whole-exome sequencing

Whole-exome sequencing was performed in collaboration with Sistemas Genómicos, S.L. (Paterna, Valencia, Spain) and the MSKCC Integrated Genomics Operation (NY, USA).

3.5.1 Library construction

DNAs from tumor samples with at least 90% of tumor cells were selected after testing their quality and integrity as previously described (section **3.3**). Library construction was perfomed using 1µg of DNA and the last version of SureSelectXT Human All Exon + UTR (71 Mb) enrichment kit (Agilent) (V4 in the ovarian cancer and V5 in the endometrial cancer projects), following manufacturer's protocols and recommendations. The quality and quantity of the libraries were analyzed using Bioanalyzer 2100 High Sensitivity assay (Agilent) and real-time PCR in LightCycler 480 (Roche), respectively.

3.5.2 Template preparation

An in-emusion PCR was performed for fragment clonal amplification, followed by an enrichment process and chemical modification that allowed sample loading in the reaction chamber. Quality and quantitiy of the microspheres from each library were estimated using the parameters from the workflow analysis.

3.5.3 Sequencing process

3.5.3.1 Ovarian cancer project

Sequencing reactions were performed to get read pairs in SOLiD high-throughput sequencing platform and a mean coverage value of 50 reads per base (50x). Parameters from SOLiD Experimental Tracking System software were used to calculate quality data (https://tools.thermofisher.com/content/sfs/brochures/cms_057557.pdf).

3.5.3.2 Endometrial cancer project

Prior to clusters generation in cbot (Illumina), a pool of the libraries was performed. Libraries were read by paired-end sequencing in an Illumina HiSeq 2500 system to a mean coverage value of 50 reads per base (50x).

3.5.4 Bioinformatics analysis

3.5.4.1 Ovarian cancer project: variant analysis

Exomes alignment to Hg19/GRCh37 human reference genome was performed using Bioscope (http://solidsoftwaretools.com). Picard-tools (http://picard.sourceforge.net) and SAMtools[200] (http://samtools.sourceforge.net/) were used for the post-alignment sequence filtering, removing that sequences with a mapping quality lower than 1 and PCR duplicates. Variant calling was performed using a combination of three different algorithms: SAMTools, GATk (http://www.broadinstitute.org/gatk/) and Bioscope. The identified Single Nucleotide Variants (SNVs) were annotated using the 'Application Programming Interfaces' (APIs) from Ensembl v64[201] and several 'custom' scripts. Biological consequence of the genomic variant in the corresponding protein was analyzed through SIFT[198], Condel[202] and Polyphen[203] computational predictors. Variants were considered mutations with negative consequence in the protein when one of the three predictors showed a damaging effect.

3.5.4.2 Ovarian cancer project: functional analysis

In order to determine the biological pathways affected by genetic variants identified in the ovarian cancer WES study, a functional analysis was performed using David protocols[204, 205]. Genetic variants were annotated using different biological databases as Biocarta (http://www.biocarta.com/), Gene Ontology (http://geneontology.org/), KEGG (http://www.genome.jp/kegg/pathway.html) and Reactome (http://www.reactome.org/). An adjust p-Value threshold of 0.2 was used for the selection of functional categories. Correlation between genes and functional categories was represented through Cytoscape tool[206].

3.5.4.3 Endometrial cancer project: variant analysis

Bioinformatics analyses were performed as previously described[207-209]. Reads from WES and targeted massive parallel sequencing were aligned to the human reference genome Hg19/GRCh37 using Burrows-Wheeler Aligner (BWA)[210]. Somatic genetic variants were identified using MuTect[211] for SNVs, and Strelka[212] and VarScan 2[213] for small insertion and deletions (indels). The potential functional effect of each non-synonymous SNV was *in silico* analyzed as previously described[209], using a combination of multiple predictors as

shown in **Figure 7**. Somatic copy number alterations (SCNAs), including loss of heterozygosity (LOH), were analyzed using TITAN[214]. The cancer cell fraction (CCF) was calculated using ABSOLUTE and manually reviewed as recommended[43, 196, 215].

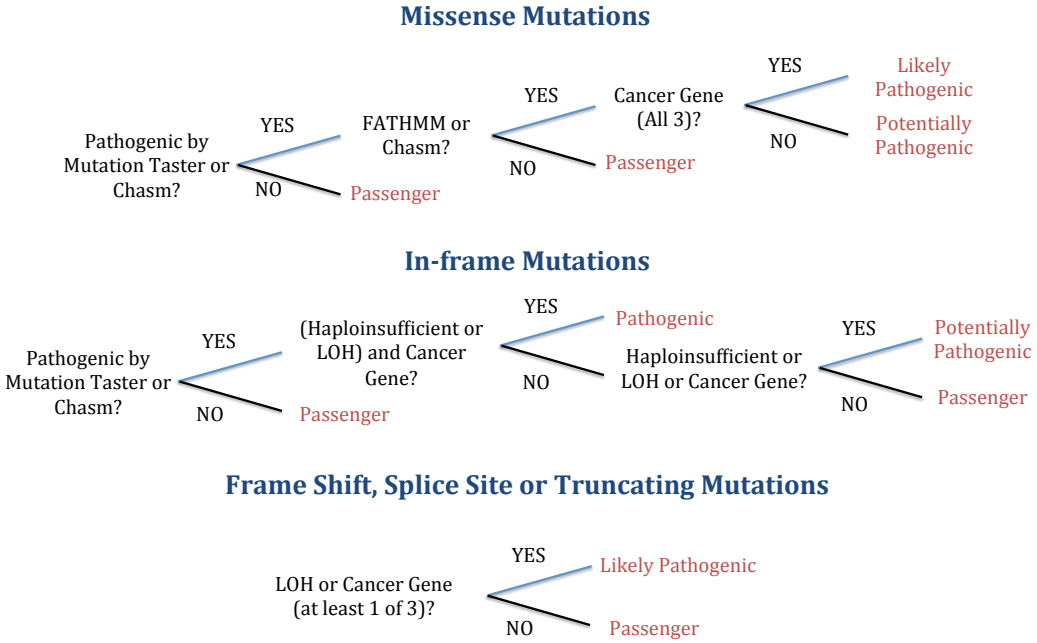


Figure 7 Decision trees followed to define pathogenicity in mutations in the endometrial cancer WES project. The possible pathogenicity of each mutation was define based on the above decision trees, depending on the type of consequence generate for each variant. According to these trees, mutations were classified as passenger, potentially pathogenic, likely pathogenic or pathogenic.

3.5.4.4 Endometrial cancer project: mutational signatures

Somatic mutations occur as a consequence of diverse processes such as mutagen exposures or defective DNA repair[216]. The specific combinations of mutation types that occur in a tumor sample are defined as mutational signature[16, 217, 218]. The mutational signature of each endometrial tumor sample included in this project was obtained by the analysis of the mutational context as previously described[16, 209]. Briefly, SNVs were categorized into C>A, C>G, C>T, T>A, T>C or T>G and then subcategorized according to the nucleotides preceding (5') and succeeding (3') the mutated base. The mutational patterns found in our samples were compared to the previously published by Alexandrov *et al*[16], after performing a normalization according to the frequency of nucleotide changes observed in the respective sequencing platforms.

3.5.4.5 Endometrial cancer project: TumorTracer Server

TumorTracer server was used to determine the origin of an specific tumor sample regarding its molecular profile[219]. Briefly, WES data was analyzed by the server in order to compare the number of nonsynonymous mutations, the mutational signature and the copy number profile with a set of 232 specific cancer genes. Public data from the COSMIC

database were used to train random-forest classifiers to discriminate among the tissue of origin. Based on these data, the classification score was defined as the proportion of trees from the validation dataset, that is, the fraction of samples with similar molecular alterations that your input sample. On the other hand, the confidence score represented the difference between the individual classification scores for the two highest-scoring tissues.

3.5.4.6 Endometrial cancer project: in silico prediction of drug treatment

In order to identified potential effective drugs according the mutational profile obtained by WES analysis, we perfomed an *in silico* study using different compounds databases: CTD (http://ctdbase.org/) and STITCH (http://stitch.embl.de/)[220-222]. These databases contain peer-reviewed information about the effect of a particular drug in a specific gene or signaling pathway, providing the correlationship between chemicals and proteins on the human context. In this analysis we included the mutational status of each identified gene in the WES study of the ambiguous endometrial case (**4.4** section). The identified drugs were used to treat patient derived xenografts (PDX) models originated from this patient (see **3.11** section). The significance was performed using the non-parametric Mann-Whitney test.

3.6 Phylogenetic analysis

3.6.1 Ovarian cancer project: phylogenetic analysis

Phylogenetic analysis was carried out using a hierarchical cluster based on mutational pattern identified in the WES study of the ovarian cancer patient. The unsupervised analysis was performed using the SPSS 17.0 statistical program (SPSS Inc., Chicago, IL) assuming Euclidean distances between mutations.

3.6.2 Endometrial cancer project: phylogenetic analysis

Maximun parsimony phylogenetic trees were generated using CCF values (in the WES analysis) or binary presence/absence matrices (in the massive targeted sequencing analysis) based on the non-synonymous somatic mutations of each tumor region, as previously described[209]. Briefly, the R package Phangorn was used to performed Neighbor-joining method and Hamming distance, optimized using the parsimony Ratchet method[223].

3.7 PCR and Sanger sequencing

A percentage of genetic variations (SNVs and INDELs) detected by WES or targeted massive parallel sequencing in the different projects were validated by Sanger sequencing as a second independent method.

3.7.1 Primer design and PCR conditions

Primers for the selected variant validation were designed using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and aligned to Hg19/GRCh37 genomic reference using Blast (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to test their specifity. Forward and reverse primers were designed to have a similar melting temperature (Tm) and a CG content between 40 and 60%. In order to stablish the optimal PCR annealing temperature for each pair of primers, a gradient PCR protocol was performed. A total of 100ng of DNA were mixed with diluted primers (10X, 2.5µl), PCR master mix (M750B, Promega) (2X, 12.5µl) and H_2O in a final volume of 25µl. The following protocol was run in a T100TM Thermal Cycle (Bio-Rad): 5min at 94°C; 40 cycles of 45s at 94°C, 45s at a range of 55-59°C and 45s at 72°C; and 7min at 72°C. PCR products were analyzed by electrophoresis using 2.5% agarose gels (Agarose D1 Low EEO, Conda-Pronadisa) stained with SybrSafe (ThermoFisher) and visualized in a Chemidoc XRS+ Molecular Image (BioRad). The annealing temperature resulting in the most efficient amplification was chosen for each pair of primers, and it is shown in Supplementary Table 6.

3.7.2 Sequencing and data anlysis

Sequencing reactions were prepared with BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems). The sequencing mix included 20 ng of DNA from the PCR product per each 100 bases of amplicon length and 1µl of primer (5pM). Products were purified using gel columns with the Optima DTR 96 Well Plates (EdgeBio), in order to eliminate the dye terminators, dNTPs and primers non-included during the reaction, to avoid they interfere in the sequencing signal. Sequencing products were analyzed in a 3730xl DNA Analyzer (96-capillary, ThermoFisher). Sequencing analyses were performed at the Genomic Unit of the Madrid Science Park.

3.8 Microsatellite instability study

Microsatellite instability (MSI) was determined in the EC samples analyzed in the WES project, following the recommended criteria[196, 224, 225]. As an initial approach, immunohistochemistry analyses of MLH1, MSH2, MSH6 and PMS2 proteins were performed (see 3.2.5 section). Loss of expression of at least one mismatch repair protein was considered to classify the case as MSI positive[224]. After that, 5 microsatellite markers including two mononucleotide (BAT25 and BAT26) and three dinucleotide (D2S123, D5S346 and D17S250) repeat markers were analyzed by PCR amplification followed by fragment analysis. PCR was performed as previously described (section 3.7.1), and PCR conditions are included in Supplementary Table 6. Fragment analysis was performed using a 3730 DNA Analyzer (Applied Biosystem) at the Genomic Unit of the Madrid Science Park. Cases with instability detected in more than one microsatellite marker were classified as MSI-High. MSI status of EC patients (Table 8) is based on results indicated in Table 11.

Table 11 MSI analysis by IHC and microsatellite markers PCR study

	IHC				Microsatellite markers				
	MLH1	MSH2	MSH6	PMS2	D2S123	D5S346	D17S250	Bat25	Bat26
EEC-1	+	+	+	+	n.a.	n.a.	n.a.	n.a.	n.a.
EEC-2	+	+	+	+	n.a.	n.a.	n.a.	n.a.	n.a.
EEC-3	+	+	+	+	n.a.	n.a.	n.a.	n.a.	n.a.
EEC-4	-	+	+	-	-	-	-	+	-
EEC-5	-	+	+	-	+	+	-	+	+
EEC-6	-	+	+	-	-	-	-	+	+
EEC-7	-	+	-	-	+	+	-	+	+

IHC: + : expression, - : loss of expression. Microsatellite markers: + : instability, - : not presence of instability, n.a. : not analyzed. Red color marks the abnormal status.

3.9 Comparative genomic hybridization (CGH)

Comparative Genomic Hybridization (CGH) was accomplished adapting the methods described by Kallioniemi *et al.*[226]. Nick translation kit (Abbot Molecular Inc) was used to label tumor and normal DNA. A total of 200ng of labeled DNA per sample was hybridized to normal female metaphase cells in the presence of 20-35mg of Cot-1 DNA for 3 days. After washes, chromosomes were counterstained with DAPI in an antifade solution (Abbot Molecular Inc). For the visualization, a Leica DM4500 epifluorescence microscope equipped with a CCD camera was used. At least 15 metaphase cells per condition were analyzed in each case using the CytoVision System with version 7.3.1 high-resolution CGH

analysis software (Leica Biosystems, UK). An average of normal cases were used as a dynamic standard reference for comparing the CGH profiles. Along the mean ratio profiles, the 99.5% confidence interval of each mean ratio profile value was compared to the corresponding 99.5% standard reference intervals. Chromosome regions with no overlap between the two intervals were considered as being altered.

3.10 Fluorescence in situ hybridization (FISH)

To perform fluorescence *in situ* hybridization the same metaphase chromosomes were prepared directly form the samples used for CGH. Slices were incubated at 90°C for 10min, washed several times with ethanol for dehydration and denatured at 75°C for 1min in the presence of a probe. Gene-specific probes for *PML* (locus 15q22) and *RARA* (locus 17q21.1, control gene) (Vysis, Downers Grove, IL) were use for detection of gene amplication. At least 100 interphase nucleis were analyzed per sample.

3.11 Patient Derived Xenograft (PDX): generation and treatment

Patient derived xenograft (PDX) were obtained from the Biomedical Research Group in Gynaecology (Vall d'Hebron Institute of Research). All animal procedures were performed according to protocols approved by the Animal Experimentation Ethics Committee from the Vall d'Hebron University Hospital. Five patient PDXs from an ambiguous endometrial carcinoma (AEC_PDX) were generated using fresh primary tumor tissue by subcutaneously implantation in the animal flanks. Regions from the superficial tumor (AEC_PDX1), deep tumor (AEC_PDX2), right lymphatic node (AEC_PDX3), left lymphatic node (AEC_PDX4) and cervix (AEC_PDX5) were implanted. After growing in the mice during around 2 months, tumors were resected and processed for formalin fixation or freezing procedure (**3.2.1** and **3.2.2** sections). DNA for sequencing analyses was extracted as previously described (**3.3** section).

In order to validated the potential effect of drugs identified in the *in silico* study (***3.5.4.6*** section), small pieces of the AEC_PDX1 model were surgically transplanted subcutaneously into female Swiss nude mice, 6 weeks old, and allowed to establish. Two mice of this second generation were analyzed per treatment condition. Mice were randomized into the following groups: treated with placebo (serum vehicle where drugs were diluted), EC standard chemotherapy -carboplatin (40mg/kg) and paclitaxel (15mg/kg)-, bortezomib (0.25mg/kg) as a single agent and the combination of bortezomib

(0.25 mg/kg) and paclitaxel (15mg/kg). Bortezomib concentration that induces a low toxicity was selected according to previous studies[227]. Chemoterapy was performed once weekly, and bortezomib treatment was performed once every three days; both of them by intra-peritoneal injection. Tumor sizes were measured twice weekly with a vernier caliper.

3.12 Statistical analysis

3.12.1 TCGA meta-analysis performed in the ovarian cancer WES project

To analyze the implication of *TP53* status in patient outcome, a meta-analysis of the TCGA genomic data was performed using the previously published data[90]. A total of 278 patients with available WES and m-RNA sequencing data were included. In order to differenciate *TP53* mutated from *TP53* null cases, those tumors with frameshift, splice site or nonsense mutations and with an mRNA level lower than the wild-type cases (z-score lower than -1) were classified as *TP53* null tumors (**Figure 8**). The z-score value was calculated as follow:

$$z = \frac{(\text{expression in tumor sample}) - (\text{mean expression in normal sample})}{(\text{standard deviation of expression in normal sample})}$$

Following these criteria, patients were subdivided into *TP53* null (n=70), mutated (n=193) and wild-type (n=15) regarding their *TP53* status. The comparison of the copy number altered genome fraction between these subgroups was performed using one-way ANOVA test and Tukey's multiple comparison test. Overall and progression free survival analyses were visualized using Kaplan-Meier method with the SPSS Statistics 17.0 software (SPSS Inc., Chicago, IL). Log-rank (Mantel-Cox) test was used to determine the significance. p-values less than 0.05 were considered statistically significant.

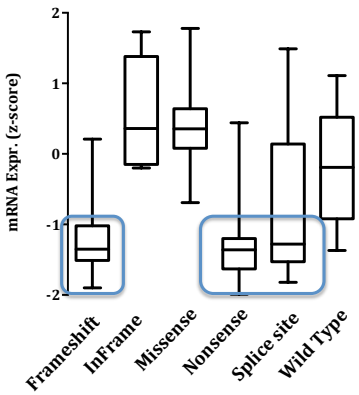


Figure 8 *TP53* mRNA expression related to mutation type in TCGA cohort. mRNA expression of the selected patients were represented regarding their *TP53* mutational status. Blue squares identify *TP53* null cases, defined as thoses cases with frameshift, splice site or nonsense mutations and with a mRNA z-score lower than -1 (lower than mRNA level in wild type cases).

3.12.2 Statistical analysis performed in the uterine aspirates project

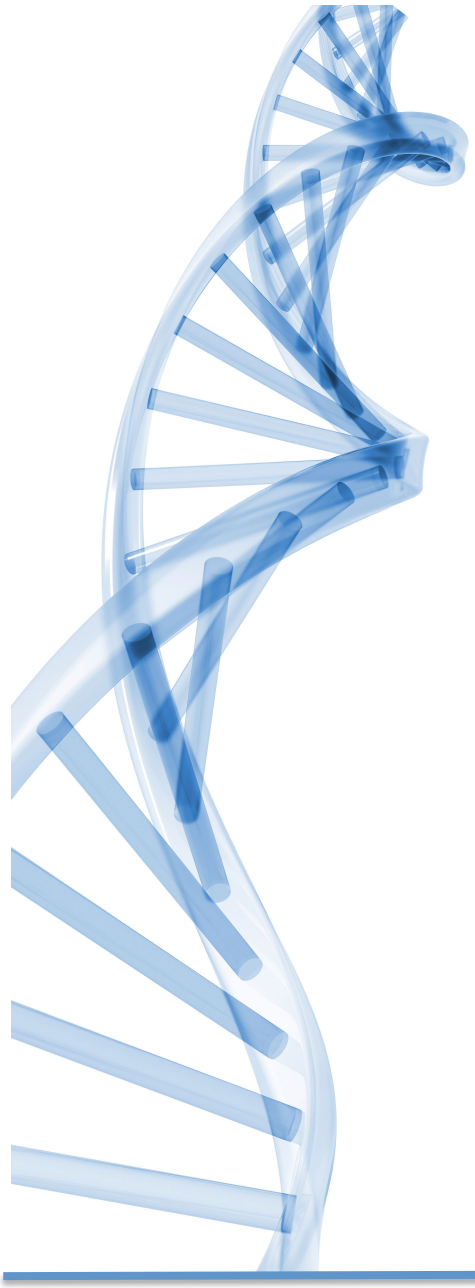
Data from hysterectomy tumor samples and uterine aspirates were compared using paired T-test. Tests were two-tailed and 95% confidence intervals were accepted. p-values less than 0.05 were considered statistically significant. Correlation between the percentages of tumor cells and Mutation Discovery Rate (MDR) values were analyzed using the Pearson coefficient. The MDR was calculated for each sample (uterine aspirate or tumor region) from the same patient according to this equation:

$$MDR = \frac{\sum Sample\ mutation}{\sum(\sum Aspirate\ mutation + \sum TR1\ mutation + \cdots + \sum TRn\ mutation)}\ x\ 100$$

Statistical analyses were performed using the SPSS Statistics 17.0 software (SPSS Inc., Chicago, IL).

3.12.3 Statistical analysis performed in the endometrial cancer WES project

Data comparison between the TCGA dataset and the patient series included in this project was performed using Mann-Withney U test. Patient subgroups were compared by unpaired T-test (comparisons between two groups) or one-way ANOVA and Tukey's multiple comparison test (comparisons between more than two groups). Statistical analyses were performed using the SPSS Statistics 17.0 software (SPSS Inc., Chicago, IL).



RESULTS

4 Results

4.1 Decoding intratumor genetic heterogeneity in a recurrent *TP53* null HGSOC patient

Intratumor heterogeneity (ITH) has been previously described in ovarian cancer and mostly in HGSOC patients[111, 113, 114], a disease characterized by a high frequency of *TP53* mutation among other features[90]. Due to the potential relation between *TP53* status and patient prognosis, a meta-analysis using the TCGA dataset was performed. Next, multiple samples from a recurrent *TP53* null HGSOC were analyzed by WES and CGH analyses, in order to further characterize ITH in this tumor subtype.

4.1.1 Meta-analysis of the TCGA cohort based on the *TP53* mutational status

There are numerous studies that correlate *TP53* mutation status with clinical prognosis, although this remains controversial due to conflicting results. While many of them provide evidences pointing to a worse prognosis for null mutations, the initial studies did not consider the biological consequences of somatic *TP53* mutations, and none of them include a broad cohort of patients[104].

Due to the availability of the TCGA dataset[90], which includes more tan 300 patients diagnosed with HGSOC, a meta-analysis was performed to compare overall and progression-free survival according to *TP53* status (see **3.12.1** section). Patients were divided into *TP53* null (n=70), mutated (n=193) and wild-type (n=15) regarding their *TP53* status. Interestingly, although the number of mutations identified in each subgroup was similar, significant differences in the fraction of copy number altered genome were found (**Figure 9a**). In fact, the null subgroup showed the highest faction of altered genome. Tukey's multiple comparison test revealed that these differences were found between the null and the wild-type subgroups. Additionally, the *TP53* null subgroup showed an intermediate clinical behavior between the wild-type and mutant groups, with significant and nearly equivalent differences between these categories as is shown in **Figure 9b**. Further studies to better characterized the differences found in the *TP53* null subgroup should be carried out.

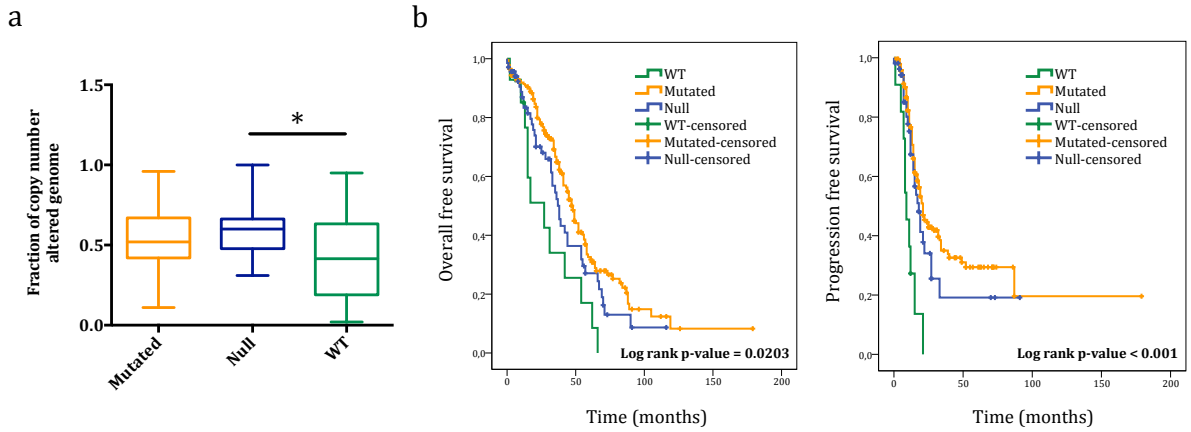


Figure 9 *TP53* mutational status meta-analysis in TCGA ovarian cancer cohort. Clinical data obtained from the TCGA study[90] were subdivided into three subgroups depending on *TP53* mutational status (see 3.12.1 section):, mutated (orange), null (blue) and wild-type (green) carcinomas. a) Comparison of the fraction of copy number altered genome between the three subgroups of tumors. b) Kaplan-Meier plots showing the association between *TP53* mutational status and overall and progression free survival. Significant correlation was performed using a Long-rank-p-value method. p-values lower than 0.05 were considered statistically significant (*p<0.05).

4.1.2 Clinical case description and sample selection

As previously mentioned, ITH has been further characterized in HGSOCs. However, none of these studies have considered the *TP53* null subgroup as an independent entity[110-118]. In order to study ITH in a *TP53* null tumor, a total of 12 samples from a 50-year-old patient with a recurrent HGSOC were analyzed (**Figure 10a**, **Table 5**). Patient was diagnosed with a stage IIIC HGSOC with extensive peritoneal carcinomatosis in 2009. The first-line treatment was based on a combination of surgical primary cytoreduction and chemotherapy. The patient received six cycles of weekly paclitaxel and carboplatin, with bevacizumab every 3 weeks for a total period of 12 months, following the OCTAVIA clinical trial criteria[228]. After a 23 months period of platinum free interval, the patient relapsed with numerous intra- and extra-pelvic implants. Because of that, the patient underwent a second debulking surgery withouth residual macroscopic disease (**Figure 10b**). The pathologist examination found papillary patterns with frequent necrosis, nuclear expression of WT1 confirming a serous histology and very high proliferative rate (90% as determined by Ki-67 staining), in both primary and relapsed tumor. Immunohistochemistry study showed lack of *PTEN* expression (suggestive of mutation) and complete absence of p53 staining (indicative of wild-type status or null mutation) (**Figure 10c**). To validate the *TP53* status of this tumor, Sanger sequencing of *TP53* coding region was perfomed, detecting a mutation in exon 7 (c.702C>G) that resulted in a premature stop codon in the position 234 of the protein (p.Tyr234*) (**Figure 10d**). These results confirmed the *TP53* null status of this tumor.

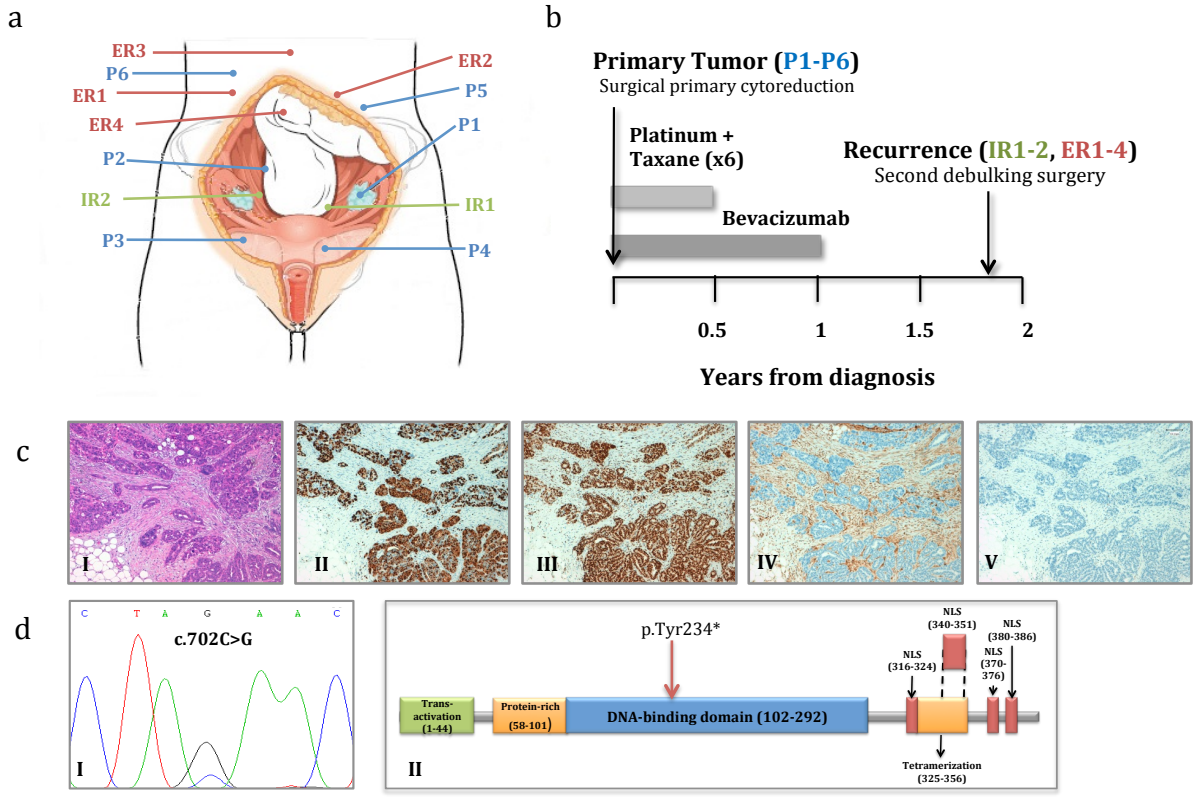


Figure 10 Sample selection and clinal case description of a *TP53* null HGSOC. a) Anatomical location of the primary tumor (P, blue) and recurrence samples (intra-pelvic, IR in green and extra-pelvic, ER in red). b) Clinical case progression. Grey squares indicate periods of treatment between primary tumor diagnosis and recurrence. c) Representative pictures of H&E staining (I) and IHC for WT1 (II), Ki-67 (III), PTEN (IV) and p53 (V). Maginification 20x. d) Electropherogram showing *TP53* mutation detected in exon 7 (c.702C>G) by Sanger sequencing (I) and diagram with the corresponding consequence in p53 protein (p.Tyr234*) (II). NLS: nuclear localization signal.

4.1.3 Whole-exome sequencing revealed differences in the mutational patterns of the distinct tumor regions

WES analysis was carried out in three frozen samples from the *TP53* null tumor, including one region from the primary tumor (P1), a pararectal recurrence implant (IR1), an ileac recurrence implant (ER1), and mesothelium as a reference normal tissue (N) (**Figure 10a**). In order to calculate the quality of the WES, the percentage of mapped reads that passed quality filters (High Quality or HQ reads) and reads 'on target' (reads aligned withing target regions) were used as standard quality metrics. All the samples reached the minimum quality criteria, with more than 70% HQ reads and more than 60% reads 'on target'. Moreover, coverage analysis showed that more than 85% of exome positions had a minimum 10x coverage (**Table 12**).

To calculate the false positive ratio of our study, a selection of 25 of the variants identified by WES were reanalyzed by Sanger sequencing. Only four of them (*ECE1*, *GTF2I*, *KMT2A*, *PCYT2*) were not detected on the validation study, revealing a false positive rate around 16% in our WES analysis (Supplementary Table 6, **Figure 11**).

Table 12 WES quality metrics from a *TP53* null HGSOC case sequencing analysis

Sample	Number of Reads	% 'HQ' Reads	% Reads 'on target'	% Exome positions with coverage >10
N	73640690	70.84	60.14	89.12
P1	73200254	70.44	60.42	90.4
IR1	62298712	71.59	61.68	91.65
ER1	71276218	70.98	61.22	93.31

Bioinformatics analysis comparing tumoral tissue with the normal control identified a total of 102 variants, 99 of them single nucleotide variants (SNVs) and three deletions (Supplementary Table 7). Only the 41% (42/102) of identified variants were present in all the samples, whereas the rest of them were specific of one or shared between two samples (**Figure 11a**). In order to analize the consequence of each variant in the corresponding protein, Condel, SIFT and Polyphen computational predictors were used[198, 202, 203]. These analyses revealed that 80 of the 102 variants (78%) showed a probably damaging or deleterious consequence in their corresponding protein (**Figure 11b**). In conclusion, our results showed a clear intratumor genetic heterogeneity in this patient, with different mutational profiles between the multiple samples.

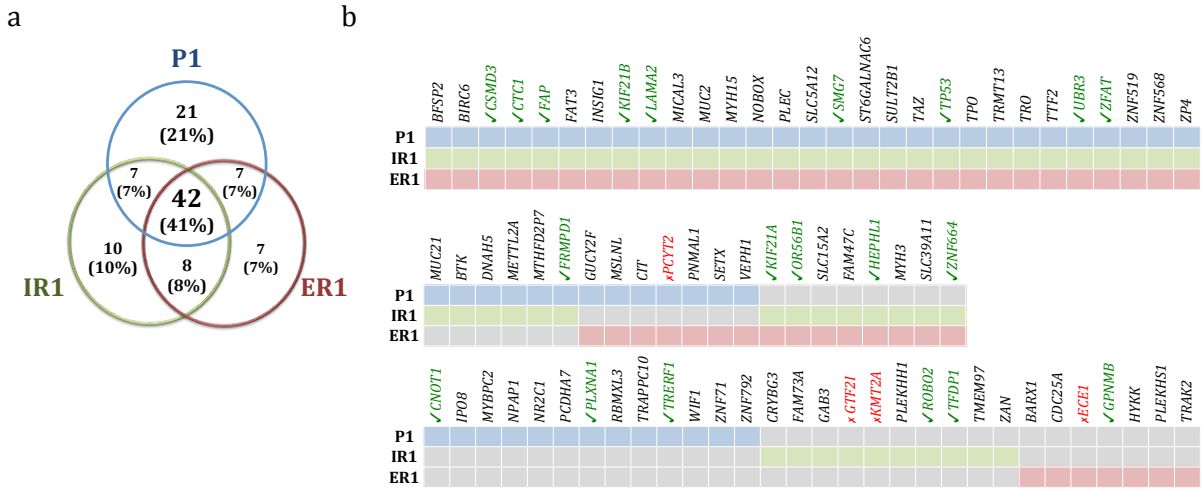


Figure 11 Variant distribution identified by WES analysis in a *TP53* null ovarian tumor. a) Venn diagram representing the number and percentage of variants among different samples (P1, primary tumor region: blue; IR1, intrapelvic recurrence sample: green; ER1, extrapelvic recurrence sample: red). b) Genes with probably pathogenic variants identified by WES. Colored squares represent presence of the mutation in the corresponding sample (P: blue, IR: green, ER: red). Light grey squares represent lack of mutation (wild-type status). Variants that were re-analyzed by Sanger sequencing are marked with the ✓ symbol (in green) if were validate, or marked with the ✗ symbol (in red) in case they were not detected.

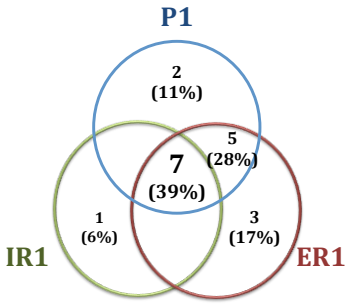
4.1.4 CGH analysis confirmed genomic ITH with an unequal distribution of SCNAs among the different tumor regions

HGSOCs frequently present high chromosomal instability and numerous somatic copy number alterations (SCNAs)[90]. To further deep on the genomic profile of this tumor, comparative genomic hybridization (CGH) was performed in the samples previously analyzed by WES (P1, IR1, ER1 and N). A total of 18 chromosomal regions with changes were identified, being more frequent the gain of material than the loss (61% vs 39%) (**Figure 12a**).

a

Enhanced		Diminished	
Region	**Sample**	**Region**	**Sample**
1p22p35	P1+ER1	-4	ALL
2q32q33	ALL	6q25q27	ALL
3q22q29	ALL	8p22p23	P1
5p15	ER1	-12	P1+ER1
7q22q32	ALL	-16	P1+ER1
8q12q24	ALL	16q24	IR1
9q31	P1+ER1	17p13	P1+ER1
10q22	ER1		
11q14q22	ALL		
15q22q26	P1		
22q13	ER1		

b



c

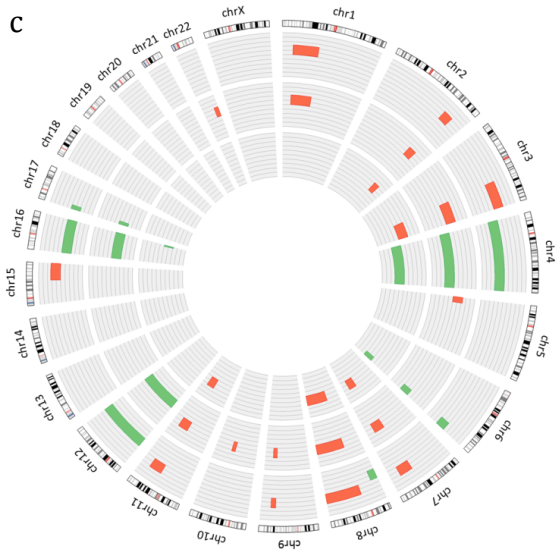


Figure 12 Somatic Copy Number Alterations (SCNAs) identified by CGH study in a *TP53* null ovarian tumor. a) SCNAs detected by CGH in samples from the *TP53* null tumor. b) Venn diagram representing the distribution of the total number of SCNAs among the different samples (P1, primary tumor region: blue; IR1, intrapelvic recurrence sample: green; ER1, extrapelvic recurrence sample: red). c) Circos diagram representing the gained (red) and lost (green) regions in the three indicated samples. Concentric circles represent P1, ER1 and IR1 samples (from the outside to the inside).

Regarding SCNAs, similar results to those previously observed in the genetic analysis were found. Only 7 out of the 18 chromosomal regions (39%) were common to all the samples. The common changes counted five gains (2q32q33, 3q22q29, 7q22q32, 8q12q24 and 11q14q22) and two lost regions (6q25q27 and depletion of whole chromosome 4). Additionally, P1 and ER1 samples shared 5 variations (gain of 1p22p35 and 9q31, and loss of 17p13 and depletion of

whole chromosomes 12 and 16), while changes common to P1 and IR1 or to IR1 and ER1 were not found. However, all samples presented exclusive alterations (8p22p23 and 15q22q26 in P1; 16q24 in IR1; and 5p15, 10q22 and 22q13 in ER1) (**Figure 12b** and **c**). Therefore, ITH was also found at the genomic level between the three analyzed samples in this patient.

4.1.5 The study of additional tumor regions allowed the identification of genetically heterogeneous tumor subclones and the phylogenetic tumor evolution

To better understand the ITH previously characterized in a *TP53* null HGSOC patient, nine additonal FFPE samples representing different regions of the primary tumor and recurrences (P2-P6, IR2, ER2-ER4) were selected for further analysis (locations are represented in **Figure 10a**). A selection of 20 previously detected variants was re-analyzed by Sanger sequencing in the four WES samples (N, P1, IR1, ER1) and in nine additional FFPE samples (**Figure 13a**).

These variants were selected to represent the different sample distribution observed in WES: nine ubiquitous (*TP53*, *CSMD3*, *CTC1*, *FAP*, *KIF21B*, *LAMA2*, *SMG7*, *UBR3* and *ZFAT*), one shared by P1 and IR1 (*FRMPD1*), four shared by IR1 and ER1 (*HEPHL1*, *KIF21A*, *OR56B1* and *ZNF664*), three P1 specific (*CNOT1*, *PLXNA1* and *TRERF1*), two IR1 specific (*ROBO2* and *TFDP1*) and one ER1 specific (*GPNMB*). Only 7 of the 20 (35%) analyzed variants were common to all the samples, reflecting the ITH previously observed in the WES results. As expected, one of these variants was the *TP53* nonsense mutation, whose ubiquitous presence is consistent with the founder role of this gene in HGSOCs. Conversely, variants in *FAP* and *ZFAT* genes, initially common to all samples according to the WES analysis, were not shared when we studied more samples by Sanger sequencing; while *FRMPD1* variant, originally detected in P1 and IR1 but not ER1 samples, was detected in the rest of the samples, appearing to be almost ubiquitous in this tumor. Interestingly, the majority of the variants previously considered as exclusive of the recurrence samples (3 of 4, 75%), were identified in three of the primary tumor regions (P3-P5), whereas only *ZNF664* variant was detected in all the recurrence samples but not in the primary tumor regions. This could be an example of genetic variant adquisition during the development of the recurrent disease. Finally, the variants previously classified as sample-specific remained being exclusive of the different main regions of the tumor (primary tumor, intrapelvic or extrapelvic recurrence) although not always restricted to an unique sample.

Taking into account that this mutational heterogeneity could be consistent with a clonal composition, a hierarchical clustering analysis based on the presence or absence of the previous studied variants was carried out (**Figure 13b**). The phylogenetic tree generated showed the existence of two main branches, with the majority of the recurrence regions clustered in the upper one along with P3, P4 and P5 samples. This branch could represent a primary tumor

ancestor subclone that gave rise to the majority of recurrence implants (all of them except ER3). In the lower branch, P1 seemed to be the closest region to the normal tissue, and P2 and P6 could be part of a different subclone from which ER3 would have arisen. These results revealed a polyphyletic evolution in this tumor, showing the importance of taking in consideration the presence of different subclones in the primary tumor that would generate the recurrence disease. Additional studies to further characterized these clones and their instrinsic featrues are necessary to better understand the implication of ITH in recurrent disease of this patient.

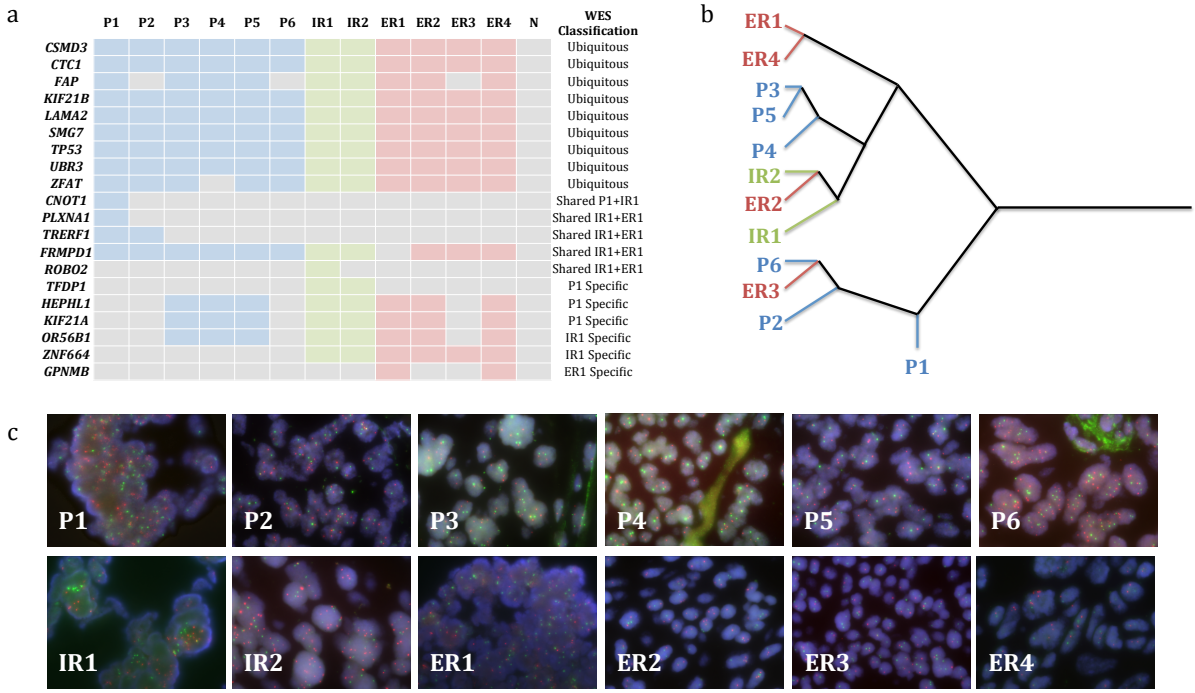


Figure 13 ITH and hierarchical clustering in additional samples from a *TP53* null ovarian tumor. a) Twenty selected variants previously detected by WES were reanalyzed in a total of 12 tumoral samples derived from the same patient. The presence or absence in each sample is represented with colors (P1-P6, primary tumor: blue; IR, intrapelvic recurrence: green; ER, extrapelvic recurrence: red) or grey squares respectively. b) Phylogenetic tree generated by hierarchical clustering based on the mutational status of each sample (see **3.6.1** section) c) Representative fluoresce *in situ* hybridization (FISH) of *PML* (red) and *RARA* (green, used as control) genes in primary tumor (P1-P6) and recurrence (IR1-IR2 and ER1-ER4) samples. Magnification 40X.

In addition to the genetic study, genomic heterogeneity found by CGH analysis (**Figure 12**) was partially validated by the fluorescence *in situ* hybridization (FISH) of *PML* (*Promyelocytic Lekemia*) gene (15q22) (**Figure 13c**). FISH analysis confirmed *PML* gain in P1 but not in IR1 and ER1 samples, in accordance with CGH data. Additional FFPE samples showed heterogeneous copy number alterations: while P3, P4, P5 and ER2 samples had 3 or 4 copies; P2, P6, IR2 and ER4 presented between 5 and 8 copies. Nonetheless, ER3 recurrence showed a potential deletion due to an unique copy of *PML* per cell was found. Therefore, ITH was also proved at a genomic level.

4.1.6 Biological functions affected by genetic variants were defined by functional annotation and network analysis

To further investigate the potential effect of genetic variants found by WES, the biological functions of the mutated genes were annotated using David protocols (see *3.5.4.2* section), through several biological databases as Biocarta, Gene Ontology, KEGG and Reactome (Supplementary Table 8). Cytoscape tool[206] was used to generate a functional network analysis, showing that cell adhesion, cell cycle control, microtubule-related movement and transport, lipid metabolism and apoptosis were pathways containing multiple mutated genes in this tumor (**Figure 14**). Additional functional analyses should be performed to better characterize the implications of these gene mutations in the corresponding pathways.

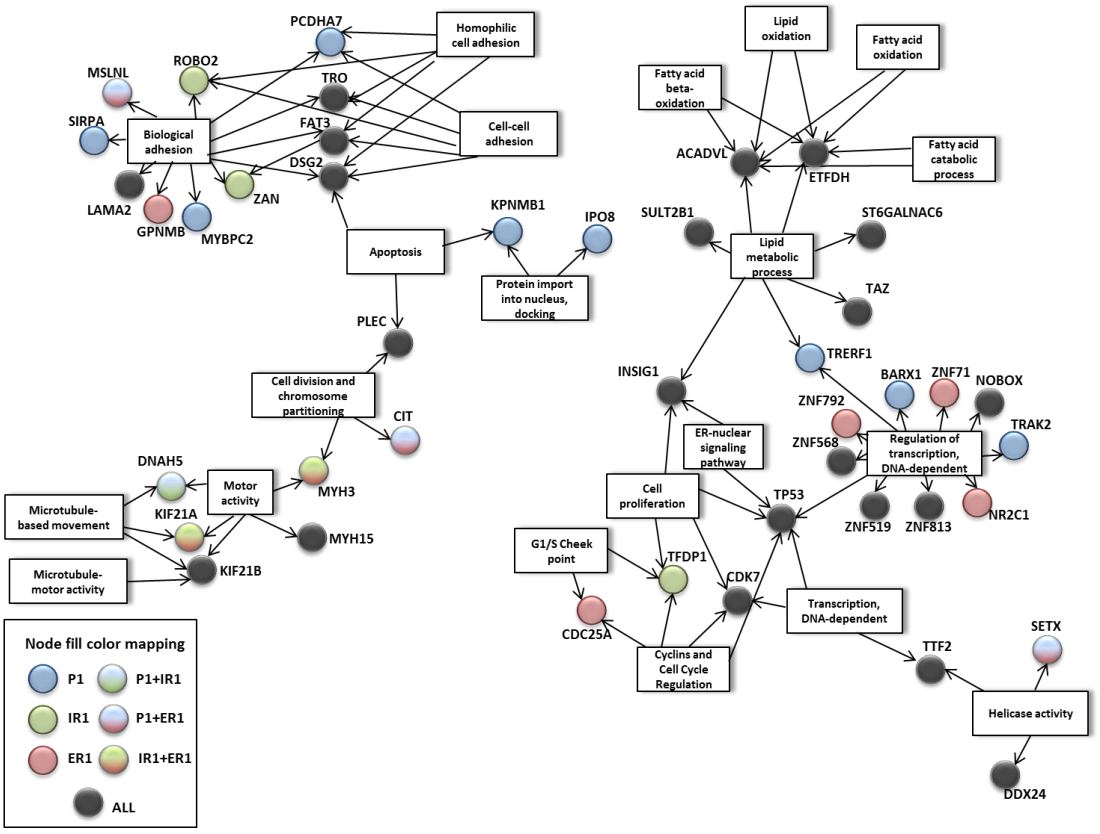


Figure 14 Functional network analysis based on annotation study revealed signaling pathways with multiple mutant genes. All the probably pathogenic variants detected by WES in a *TP53* null tumor were considered for functional annotation using David protocols and network analysis by Cytoscape tool (see *3.5.4.2* section). Results revealed signaling pathways containing multiple mutated genes, that were annotated by Biocarta, Gene Ontology, KEGG or Reactome databases and represented as squared nodes. Mutated genes were characterized as colored circles indicating the sample or samples in which each one was mutated (see node fill color mapping).

4.2 Uterine aspirates as a diagnostic tool: capturing ITH in endometrial carcinoma

Uterine aspirates (UAs) are pre-operative biopsies usually obtained with a pipelle de Cornier, which suppose a low invasive and highly sensitive procedure[149]. The use of UAs as a molecular diagnosis tool was proposed on the basis of differential gene expression profiles[153, 154]. In this thesis, the potential use of UAs to describe the tumor mutational profile before surgery and its implication in ITH have been studied. For this purpose, targeted sequencing was performed using the commercial Ion AmpliSeq Cancer Hotspot Panel v2 that analyze above 2,800 cancer mutations of 50 oncogenes and tumor suppressor genes, including some of the most frequently altered genes in endometrial cancer (*PTEN*, *KRAS*, *FGFR2*, *CTNNB1*, *PIK3CA*, *FBXW7* and *TP53*) (see **3.4.1** section). Variants were annotated as potential mutations based on PROVEAN analysis[197] (Supplementary Table 9), as previously described in the **3.4.4** section. No differences were found in sequencing quality metrics, neither between frozen or FFPE samples, nor between paired primary tumor and UAs samples (Supplementary Table 10).

4.2.1 Genetic analysis of uterine aspirates allowed tumor mutation detection

In order to determine whether the biological material from UAs was valid for the detection of genetic variants, retrospective paired samples of UAs and hysterectomy specimen were analyzed by targeted sequencing using the Cancer Hotspot Panel as mentioned above. Initially, 28 paired-samples from malignant disorders and 27 UAs from patients with bening lesions or normal endometrium were analyzed to determine their mutational status. Malignant disorders included 21 tumoral samples (7 endometrioid endometrial carcinomas, EECs; 7 serous endometrial carcinomas, SECs; and 7 carcinosarcomas, CSs) and 7 samples from atypical hyperplasia (AH) as a precursor lesion. The 27 control samples comprised patients diagnosed with non-atypical hyperplasia (NAH) (n=7) or leiomyomas (LM) (n=7), and 13 normal endometria (NE) (**Table 6**). In all of NE samples and in 3 of the 7 NAH cases, only UAs were analyzed due to these patients did not undergo surgery. Genetic variants were found in 20 of the 21 (95%) UAs from tumoral samples, according to the results obtained in the paired hysterectomy specimens. Only in one of the UAs derived from a SEC, we did not observed the mutations found in its corresponding hysterectomy sample. Additionally, we observed that 4 of the 7 (57%) UAs from the AH precursor lesions showed mutations, although only in 2 cases these mutations were also detected in their paired hysterectomy specimen. Conversely, only 1 of the 27 (4%) control samples showed a genetic variant (**Figure 15**). Since the UAs that presented a genetic variant came from a NE case, we did not have the possibility to compare it with its corresponding hysterectomy specimen.

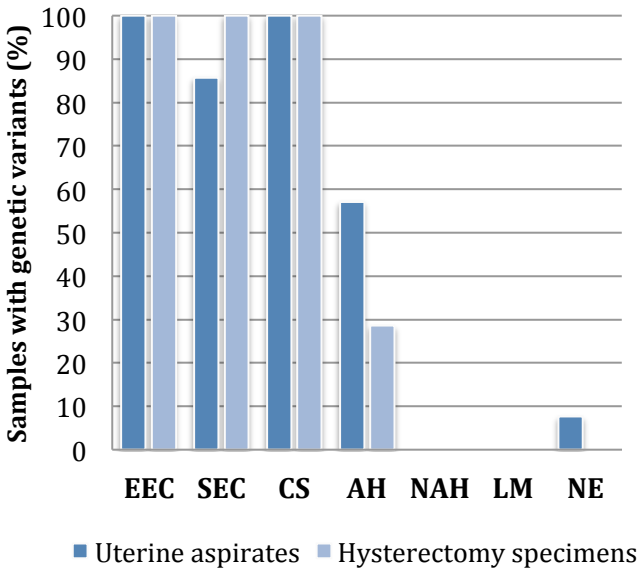


Figure 15 Percentage of samples with genetic variants detected by targeted massive parallel sequencing in paired endometrial samples. Percentage of samples in which genetic variants were identified in the targeted sequencing analysis. Paired samples of uterine aspirates (dark blue) and hysterectomy specimens (light blue) were included in the study. EEC: endometrioid endometrial carcinoma, SEC: serous endometrial carcinoma, CS: carcinosarcoma, AH: atypical hyperplasia, NAH: non-atypical hyperplasia, LM: leiomyoma, NE: normal endometrium.

4.2.2 Uterine aspirates reproduced the mutational profile of its paired hysterectomy specimen samples

Once the capability to detect mutations in UAs was demonstrated, we estimated the potential of aspirates to characterize EC from a genetic point of view. To this end, we analyzed the percentage of pathogenic variants present in hysterectomy specimen that were also identified in the UA. A total of 62 patients diagnosed with EC (44 EECs, 9 SECs and 9 CSs) and 10 patients diagnoses with AH were studied, including the previously analyzed cases. Histopathological data of these cases is detailed in **Table 7**. In this extended series, sequencing analysis revealed the presence of mutations in 58 of the 62 (94%) aspirates from cancer patients (Supplementary Table 11) and in 5 of the 10 (50%) aspirates from AH cases (Supplementary Table 12). Interestingly, in 35 out of the 44 (80%) UAs from EEC patients we identified all the mutations detected in their paired surgical tumor tissue (**Figure 16**). In fact, in only one of the EEC patients (EEC-19), the genetic analysis of UAs did not show any of the mutations preoviusly found in its hysterectomy specimens. Conversely, in one other patient (EEC-21) mutations were detected in the UA, whereas none of them were identified in its surgical sample (Supplementary Table 11). Furthermore, in 6 of the 9 (67%) SEC and in 7 of the 9 (78%) CS all the mutations identified in the hysterectomy were also detected in the corresponding aspirate (**Figure 16**). Therefore, these results confirmed that the genetic analysis of UAs reliably reproduce the genetic status of the tumor in a pre-operative setting, at least in the analyzed samples.

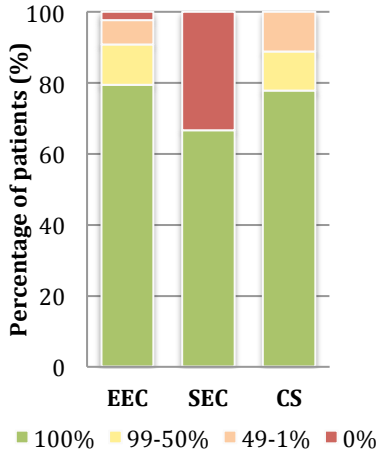


Figure 16 Percentage of hysterectomy mutations identified in paired uterine aspirates. Percentage of tumor mutations found in surgical tumor samples also detected in UAs are represented for each hystology subgroup of patients. EEC: endometrioid endometrial carcinoma. SEC: serous endometrial carcinoma. CS: carcinosarcoma

Mutations identified in the different histological subgroups of patients were consistent with previously published results[196]. Briefly, EEC carried mutations in *PTEN* (71% of the patients), *PIK3CA* (39%), *CTNNB1* (27%), *TP53* (27%), *FGFR2* (23%) and *KRAS* (21%). In addition, we also detected mutations in less commonly altered genes (**Figure 17**). As expected, the most frequently mutated gene in SEC and CS samples was *TP53* (78% and 89% respectively). Frequencies found in our series were generally higher than those previously detected in the TCGA series[196]. These differences could be explained due to the use of different sequencing methods. Thus, while TCGA study performed WES (mean coverage around to 50x), our study was developed with targeted sequencing allowing a high number of lectures (mean coverage around to 1000x) and consequently a more accurate detection of mutations[229].

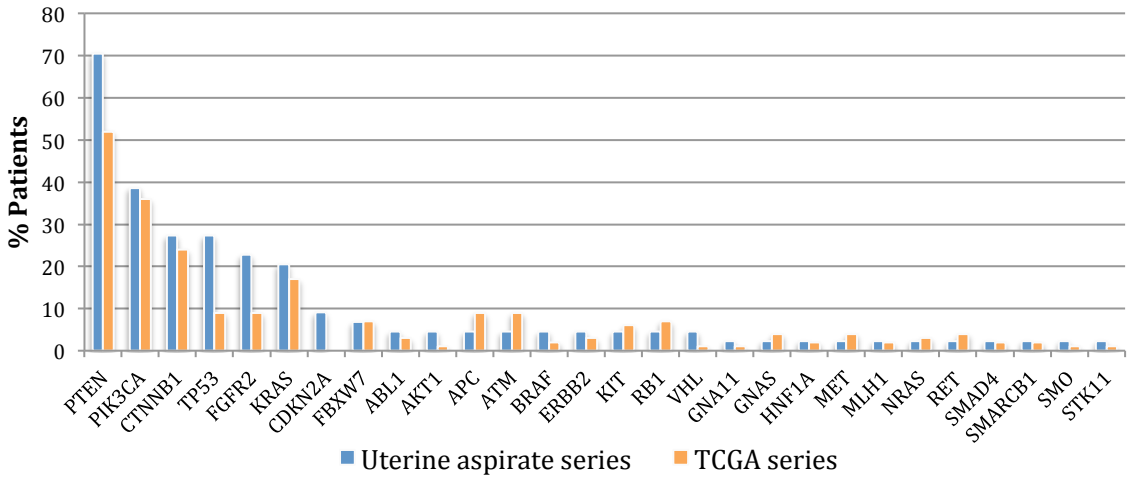


Figure 17 Mutated genes found in the endometrioid endometrial carcinomas comparing with the TCGA series. Percentage of EEC patients with mutations in the indicated genes. Comparison between our analyzed series (blue) and the TCGA dataset (orange).

4.2.3 Targeted sequencing helped to reduce diagnosis fail rate of uterine aspirates

The fail rate for pathologic analysis of UAs distinguishing the presence or absence of malignancy is around 13%, either due to a small proportion of representative tumor cells or to the poor quality of the sample[146-148]. In fact, 8 of the 62 (13%) EC aspirates analyzed in our series were considered non-diagnosable but turned out to have EC on the corresponding hysterectomy specimens (EEC38-44 and CS-9). However, genetic analysis of these aspirates revealed the presence of mutations in 7 of the 8 (88%) non-evaluable UAs and their paired hysterectomy samples, showing a similar mutational profile (Supplementary Table 11). We did not find any mutations in the UA from one of these patients as well as in its paired hysterectomy sample. These results indicated that genetic sequencing of UAs provides valuable information and could be used as a complement for pathological analysis in order to reduce its fail rate.

4.2.4 Genetic analysis of uterine aspirates captured the ITH found in ECs

Comparative mutational profile between UAs and hysterectomy specimens highlighted the presence of additional mutations, not detected in their paired surgical samples, in 11 out of the 62 (18%) aspirates. These results could be reflecting the presence of ITH in this clinical context. To explore this hypothesis, additional tumor regions from 24 of the EC previously studied (EEC1-14, SEC1-5 and CS1-5) were analyzed by targeted sequencing (**Table 13**).

Table 13 Endometrial cancer cases studied by targeted massive parallel sequencing in the ITH analysis

Patient	Tumor regions analyzed	Total variants detected in tumor regions	Total variants detected in uterine aspirates	Intratumor heterogeneity
EEC-1	3	5	5	Yes
EEC-2	2	3	3	Yes
EEC-3	3	4	4	Yes
EEC-4	4	2	5	Yes
EEC-5	3	2	2	Yes
EEC-6	3	3	4	Yes
EEC-7	4	9	4	Yes
EEC-8	4	2	2	No
EEC-9	4	4	4	Yes
EEC-10	4	1	1	No
EEC-11	4	2	2	No
EEC-12	4	10	9	Yes
EEC-13	3	12	16	Yes
EEC-14	4	3	3	No
SEC-1	3	1	1	No
SEC-2	3	1	1	No
SEC-3	3	1	1	No
SEC-4	3	1	1	Yes
SEC-5	3	4	4	No
CS-1	4	4	4	Yes
CS-2	3	1	1	No
CS-3	3	1	1	No
CS-4	3	1	1	No
CS-5	2	1	1	No

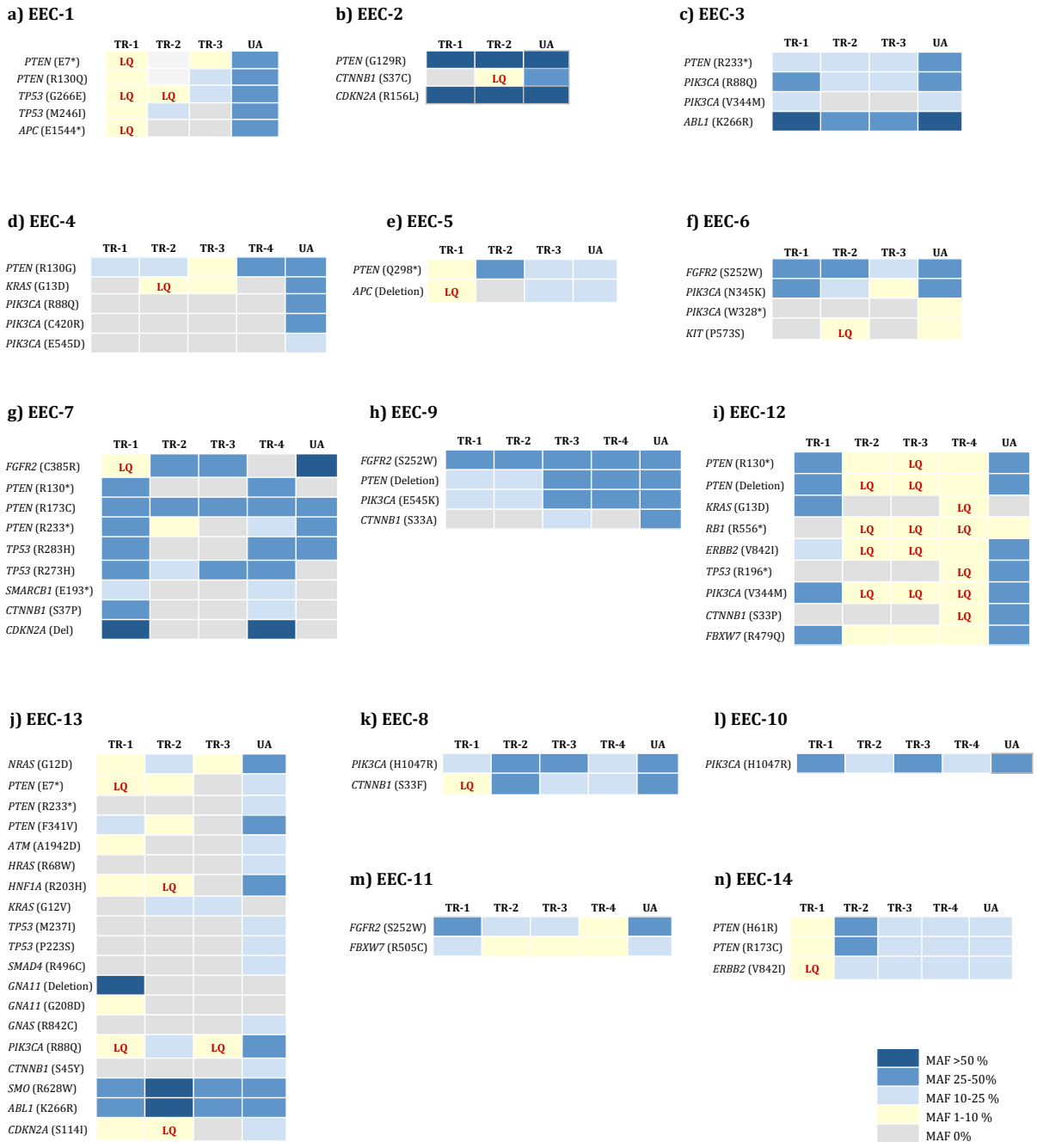


Figure 18 Intratumor heterogeneity analyses in endometrioid endometrial carcinoma samples. Results obtained in targeted massive parallel sequencing analyses of multiple samples derived from hysterectomy tumor regions (TRs) and paired uterine aspirates (UAs) from 14 endometrioid endometrial carcinoma (EEC) patients identified genetic heterogeneous (a-j) and non-heterogeneous (k-n) cases. Colors represent mutant allel frequencies (MAFs). Squares marked with LQ identified Low Quality variants detect in the Ion PGM analysis.

The comparison of the mutational profile in additional hysterectomy regions of EECs revealed that 10 of the 14 patients (71%) showed different mutations in the several tumor samples analyzed, confirming the presence of ITH (**Figure 18a-j**). This is the case of EEC-1 patient, in which three different tumor regions (TR) were analyzed. Five mutations were detected in *PTEN*, *TP53* and *APC* genes in TR-1, whereas in the other two tumor regions (TR-2 and -3) we only

observed 2 and 3 mutations respectively. Interestingly, these five mutations were identified in the paired UA, with higher values of mutant allele frequency (MAF) and better sequencing quality (**Figure 18a**).

On the other hand, 4 of the 14 EEC cases showed a similar mutational profile in all the samples analyzed, although with different MAF values (**Figure 18k-n**). For example, mutations in *PIK3CA* and *CTNNB1* were detected in the four tumor regions and the corresponding aspirate of the EEC-8 patient (**Figure 18k**). These results could be suggesting that these cases did not harbor significant ITH, at least in the sequenced tumor regions.

Contrary to EEC, ITH was only observed in 1 of the 5 (20%) SEC and 1 of the 5 (20%) CS (**Figure 19**). This low proportion of heterogeneous cases could be due to the fact that chromosomal instability is the hallmark of these tumors[230], in which single mutations are not so common.

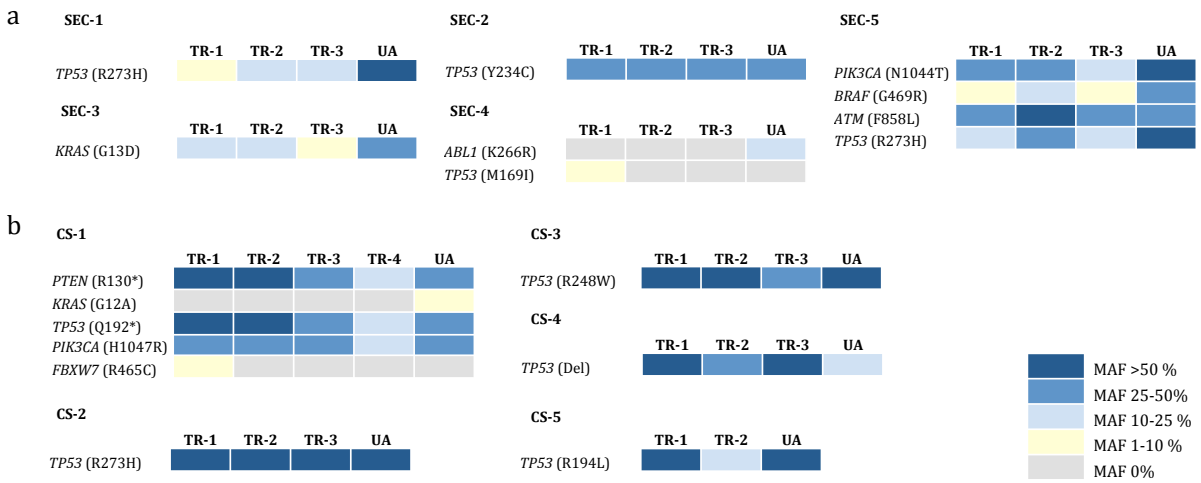


Figure 19 Intratumor heterogeneity analysis in non-endometrioid endometrial carcinoma samples. Results obtained by targeted massive parallel sequencing of multiple tumor regions (TRs) and paired uterine aspirates (UAs) from a) 5 serous carcinomas (SEC) and b) 5 carcinosarcoma (CS) patients. Colors represent mutant allel frequency (MAF).

In order to define the sensitivity of mutation detection for each sample, the mutation discovery rate (MDR) was calculated (**Figure 20**). MDR indicates the proportion of mutations detected in each sample with respect to the total amount of mutations observed in all samples analyzed for a given patient (see **3.12.2** section). MDR value was significantly higher in EEC aspirates than in their corresponding hysterectomy specimens, with a mean of 94% for UAs and 77% for individual tumor region samples. These differences were accentuated when low-quality (LQ) mutations were not considered, decreasing MDR value in surgical tumor samples to 68% and remaining unaltered in UAs (**Figure 20a**). As expected, no significant differences were found in SEC or CS samples, since are less genetic heterogeneous tumors, at least at the analyzed genetic level. In fact, the main differences in MDR values were found in heterogeneous tumors due to

the variability in the mutational profile of each sample. In 8 out of the 10 (80%) EEC heterogeneous tumors, UAs showed a higher MDR value than that tumor region used for the pathological diagnosis (TR-1) (**Figure 20b**). Only in EEC-7 patient, MDR value was higher for TR-1 sample than for the corresponding aspirate. Interestingly, UAs had a higher MDR than at least one tumor region in all the heterogeneous EEC cases. These differences were not related to the proportion of tumor tissue contained in each analyzed region, as there was no significant correlation according to Pearson test results (r = 0.146). These results confirmed that genetic analysis of UAs detects a more representative mutational landscape of the tumor, reproducing in a single sample the ITH found in the multiple TRs.

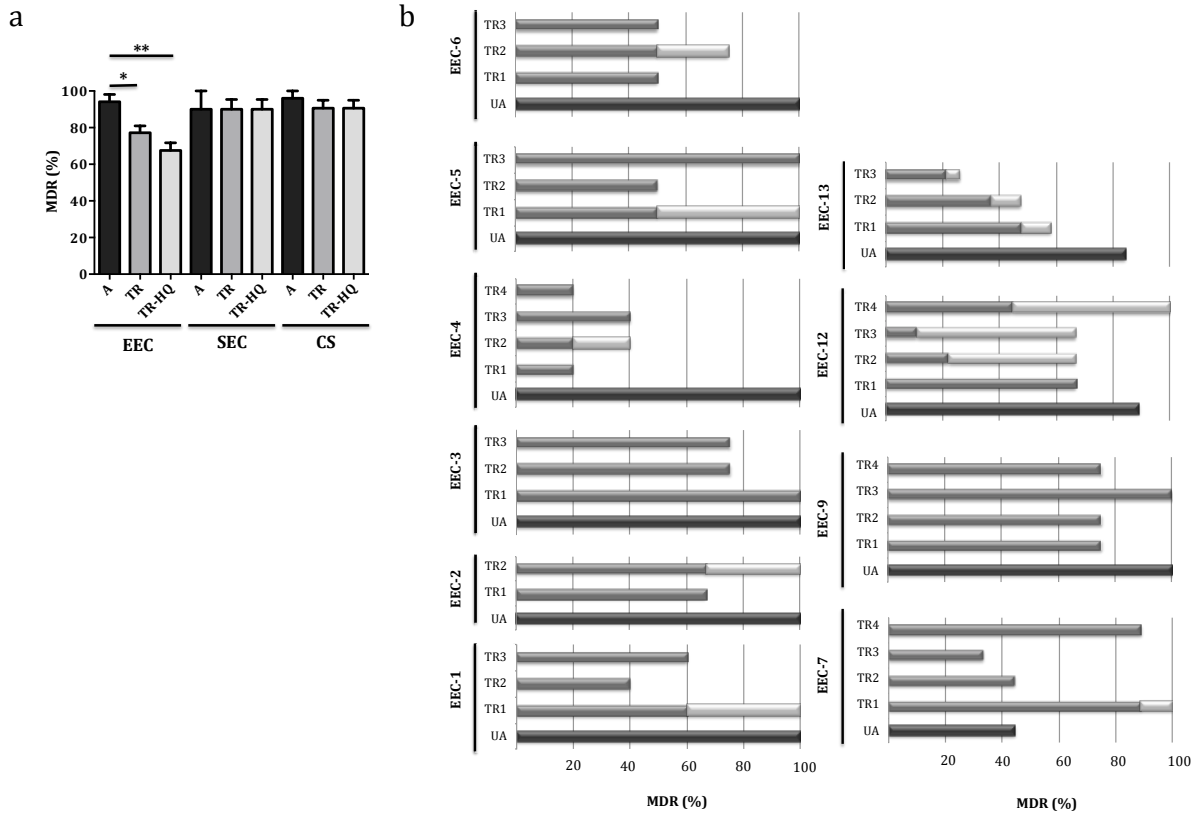


Figure 20 Mutation discovery rate (MDR) in uterine aspirates and their respective hysterectomy specimens. Mutation discovery rate (MDR) is defined as the percentage of mutations detected in each studied sample with respect to the totality of mutations observed in all analyzed samples for the same patient (see **3.12.2** section). a) MDR mean in EEC, SEC and CS patients. b) Comparison of MDR value calculated in each sample of heterogeneous EEC patients. Dark greys represents percentage of variants detected with good quality and light grey percentage of variants identified with low quality in the Ion PGM analysis. A: uterine aspirate, TR: tumor region, TR-HQ: tumor region considering high quality variants, EEC: Endometrioid Endometrial Carcinoma, SEC: Serous Endometrial Carcinoma, CS: Carcinosarcoma. (*$p<0.05$, **$0.05<p<0.001$).

4.3 Deciphering intratumor heterogeneity in metastatic endometrial carcinomas

Previous results derived from the UAs project suggested that ITH seemed to be a common event in EECs, but not in SECs or CSs (see section **4.2**). However, this study showed several limitations, considering that targeted sequencing is restricted to specific regions and only primary tumors were included. To further investigate the implication of ITH in EC, a collection of metastatic and recurrent cases was analyzed by WES. In this sense, the current study allowed not only the analysis of genetic changes as point mutations and small insertion and deletions (indels), but also big genomic variations as somatic copy number alterations (SCNAs), more common in SECs. Additionally, WES analysis was followed by a targeted sequencing validation process, which let us to increase the number of samples analyzed in each patient.

4.3.1 Cases description and endometrial samples selection

A series of 11 patients diagnosed with metastatic EC and complete follow-up information were included in the study, counting 7 EEC, 3 SEC and one ambiguous carcinoma (AEC) (**Table 8**). As previously mentioned, AECs comprise those tumors with overlapping characteristics between EECs and NEECs, difficulting a clear histologycal classificationus[128, 129]. The ultramutated subgroup of EC related to *POLE* mutations was excluded, since it could hamper the analysis of ITH. The the most frecuent *POLE* mutations in EC were analyzed by Sanger sequencing, accounting for about 90% of the mutations previously described in endometrial carcinomas[231] (Supplementary Table 13). Furthermore, microsatellite instability (MSI) was analyzed in the EEC cases by IHC and PCR amplification of 5 microsatellite markers (see **3.8** section). According to these results 4 of the 7 EEC showed MSI (EEC4-7), being 3 of them classified as MSI-high cases (EEC5-7). After sample classification, the fresh-frozen tumor tissue and the corresponding germline DNA were analyzed by WES, including at least two tumor regions from the primary tumor and one from the metastasis. Then, the validation of the identified variants was performed by targeted ampliseq sequencing, including additional FFPE samples from each case to further explore ITH in these patients (**Table 9**).

4.3.2 Whole-exome sequencing study and molecular classification of metastatic EC samples

A total of 48 frozen samples from the 11 patients (with a mean of 4 samples per patient) were analyzed by WES to a 51x median depth (range = 32-74x). Bioinformatics analysis identified a median of 83 variants per sample (range = 23-1000), 57 of them non-synonymous mutations (range = 17-779) (**Figure 21a**, Supplementary Table 14), similar to the TCGA dataset (TCGA

patients excluding *POLE* mutated subtype; median = 50, range = 14- 3810, Mann-Whitney U test p= 0.76).

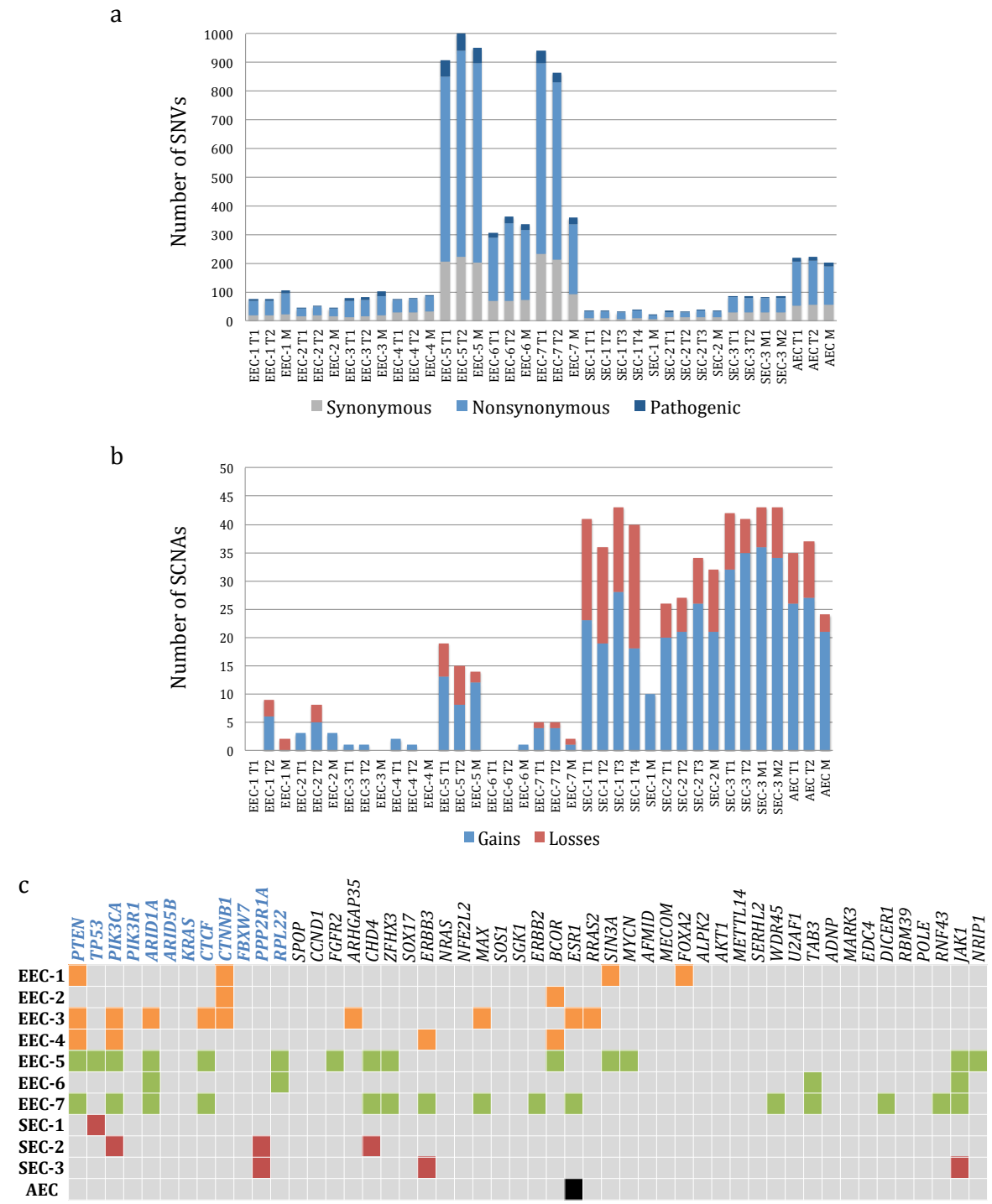


Figure 21 Summary of WES results from the multiple sample analysis from metastatic ECs. a) Number of synonymous, non-synonymous and pathogenic variants and b) number of somatic copy number alterations identified by WES analysis in specified samples. c) Frequently mutated genes previously described in endometrial carcinomas[50] (blue bold font indicates those initially described by the TCGA consortium[196]). Colored squares represent the presence of mutation in at least one sample of the patient. CN-low cases (EEC 1-4) are marked in orange, MSI cases (EEC 5-7) in green, serous-like cases (SEC 1-3) in red and the ambiguous case (AEC) in black. Grey squares indicate the absence of mutation (wild-type status).

The number of SCNAs widely varies depending on the histological subtype (median = 10, range = 0-43), being significantly higher in the serous cancer patients and in the ambiguous carcinoma case than in the endometrioid tumors (P<0.0001) (**Figure 21b**, Supplementary Table 15). No significant differences were found between the number of somatic variants neither SCNAs between primary tumors and metastasis. According to these results, tumors were stratified based on the TCGA classification[196] as copy-number low/endometrioid (EEC1-4), MSI/hypermutated (EEC5-7), and copy-number high/serous-like (SEC1-3 and AEC) (**Table 8**). At least one of the most frequently mutated genes described by the TCGA analysis[196] was found altered in all the cases, except in the ambiguous case that presented a missense non-pathogenic mutation in *ESR1*, recently described as a frequently mutated gene in EC[50] (**Figure 21c**).

Furthermore, the mutational process of each tumor region was analyzed (see *3.5.4.4* section), and mutational signatures were defined for each sample according to the previously described by Alexandrov *et al*[16] (Supplementary Table 16). While the majority of CN-low samples were related to signature 1, associated to the age of the patient, MSI tumors showed the signature 6, linked to DNA MMR deficiency. For the majority of the patients, the different samples showed concordant mutational signatures, except for EEC-4 and SEC-3 patients. In both cases, primary tumor samples showed signature 1 while metastatic regions had the mutational signature 6. Nevertheless, metastatic samples did not have the correct associated indels to be properly classified as signature 6, so this should not be considered as a reliable signature assignation. Interestingly, all the samples analyzed from the AEC case showed the mutational signature 13, which has not been previously described in ECs. Because this case showed unusual histological and molecular characteristics, it was analyzed separately from the rest of cases and will be discussed in section **4.4**.

4.3.3 Whole-exome sequencing analysis revealed different degrees of ITH in metastatic ECs

The comparison of the genetic variants found in the different samples from each patient revealed a mean of 44% shared mutations (range = 2-74%) (**Figure 22a**), with the MSI patients showing a significantly lower percentage (mean = 12%, p=0.0085). On the other hand, a mean of 54% (range = 4-100%) of pathogenic mutations were common between all the samples of each patient (**Figure 22b**), remaining the lower percentages those observed in the MSI cases. Interestingly, the percentage of common SNCAs was very low in all the subtypes, with a mean of 6% (range = 0-37%) (**Figure 22c**). However, the percentage of common genes affected by SCNAs was significantly higher (mean = 21%, range = 0-975; p=0.048) (**Figure 22d**), especially in the EEC-2 case that shared 97% of the affected genes but only 38% of SCNAs.

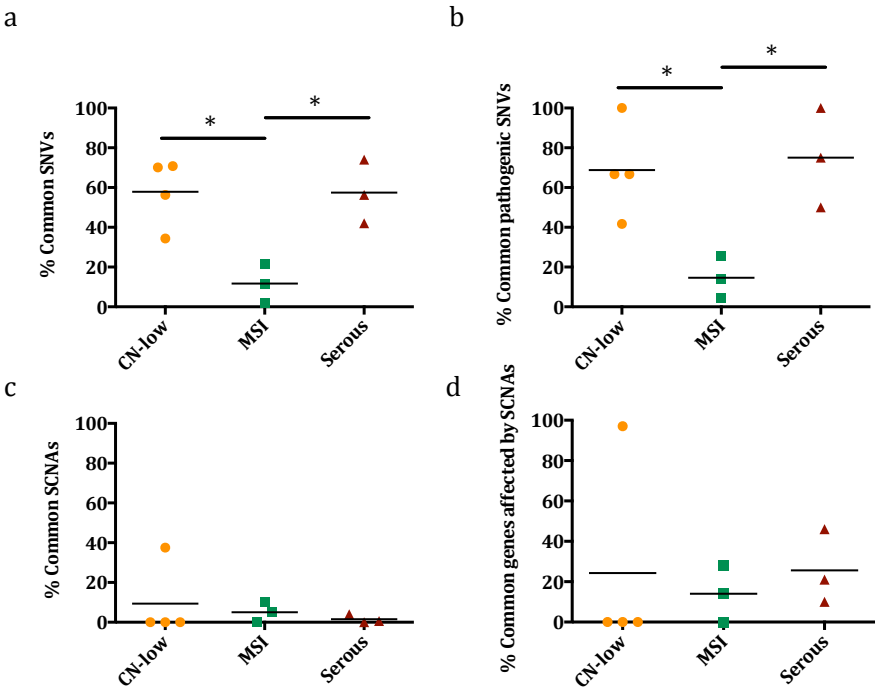


Figure 22 Percentage of common genetic and genomic alterations found in the multiple sample analysis of metastatic ECs. Percentage of shared single nucleotide variants (SNVs) (a), pathogenic SNVs (b), somatic copy number alterations (SCNAs) (c) and genes affected by SCNAs (d) identified in the comparison between multiple samples in each metastatic EC. Each point represents a patient. CN-low cases are marked in orange, MSI cases in green and serous-like cases in red. p-values lower than 0.05 were considered statistically significant (*p<0.05).

These results suggested the presence of not only genetic but also genomic ITH between the multiple samples derived from metastatic EC, with different degrees depending on the molecular subgroup. Previous results obtained in the uterine aspirates project suggested that ITH was mainly found in the endometrioid histology (see section **4.2**), but the present WES analysis revealed that ITH was also found in the serous tumors. However, this heterogeneity was more widespread at the SCNA level. Since numerous changes at the copy number level are the hallmark of serous tumors, we thoroughly investigated its distribution in these cases (**Figure 23**). While most of the genetic variants were shared by all the samples, the genes affected by SCNAs had a diverse distribution. In the SEC-1 patient, whereas 43% of the genetic variants were shared by the five analyzed samples, only 10% of the genes affected by SCNAs were common between them. Nevertheless, the percentage of common variants between the primary tumor samples was similar: 15% in the SNVs versus 20% in the SCNAs analysis. About the SEC-2 patient analysis, it is worth mentioning the high number of SCNAs detected exclusively in the metastatic region (28% of the total SCNAs identified). Finally, the SEC-3 patient also showed a higher number of shared genetic variants (73%) compared to the common SCNAs (46%). In this case, an important percentage of the SCNAs were also exclusively

observed in the metastatic regions (50/253, 20%), not being found in the primary tumor samples analyzed.

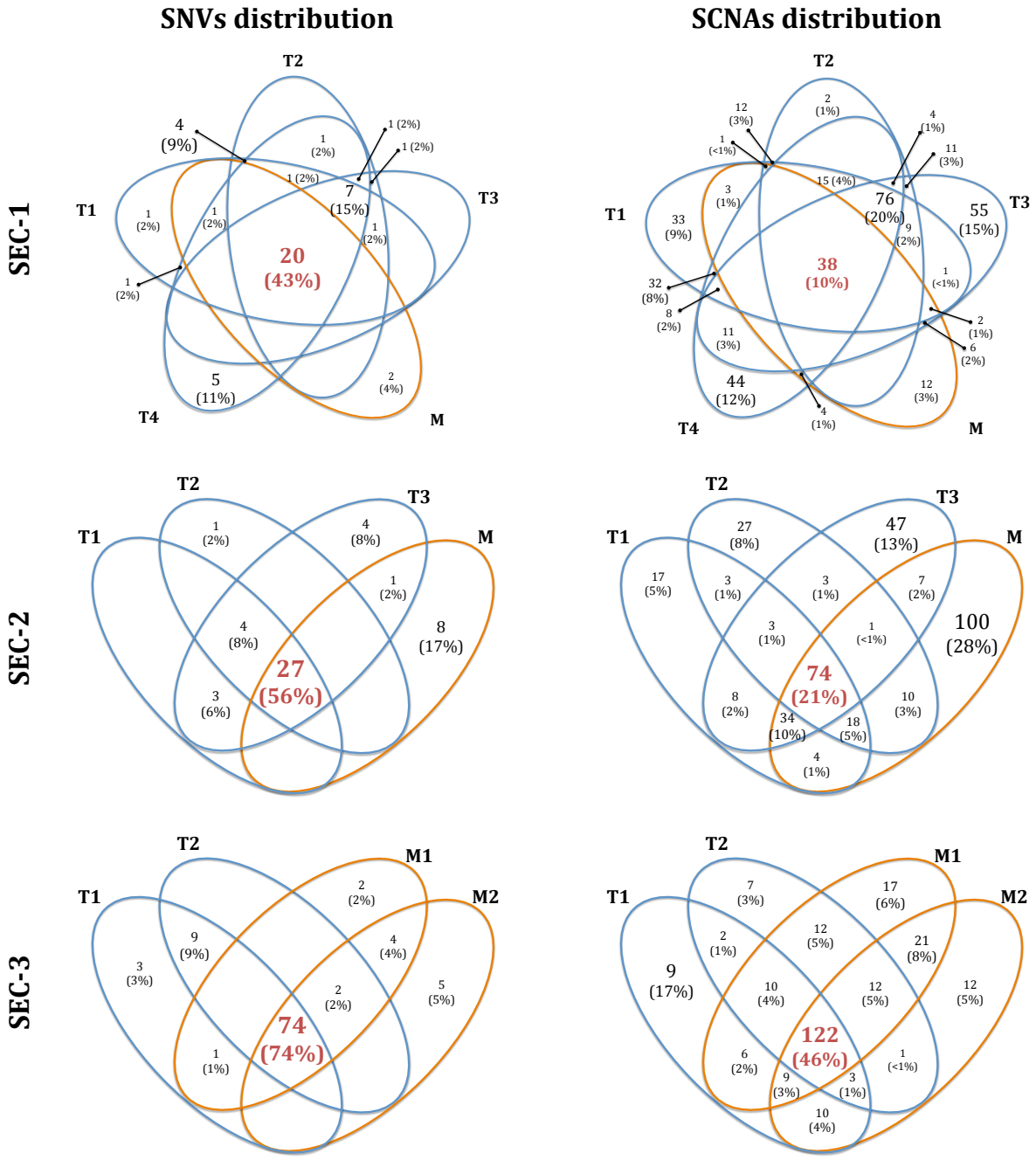


Figure 23 Genetic and genomic changes distribution in the multiple sample analysis of metastatic SECs. Venn diagrams representing the number and percentage of single nucleotide variants (SNVs, left panel) and genes affected by somatic copy number alterations (SNCAs, right panel) in the multiple samples analyzed by WES from the metastatic serous endometrial carcinomas (SECs).

To further characterize genetic ITH in each sample, tumor purity and malignant cell ploidy were inferred by the calculation of the cancer cell fraction (CCF), as previously described (see section *3.5.4.3*). CCF represents the fractions of analyzed cancer cells that carried each mutation, which

allows to discriminate between clonal and subclonal mutations in each specific sample. The majority of the detected variants revealed a clonal status (mean = 82%, range = 61-93%) (Supplementary Table 14), independently of the histological or molecular subtype of the tumor or the anatomical location of the sample (primary tumor or metastasis). In this sense, CCF values showed that most of the cancer cells of each sample carried the majority of mutations detected by WES in each specific region, whereas a small percentage of variants were only present in subclonal populations of the sample. Variants distribution and clonal status are represented in the Supplementary Figure 1 and Supplementary Figure 2.

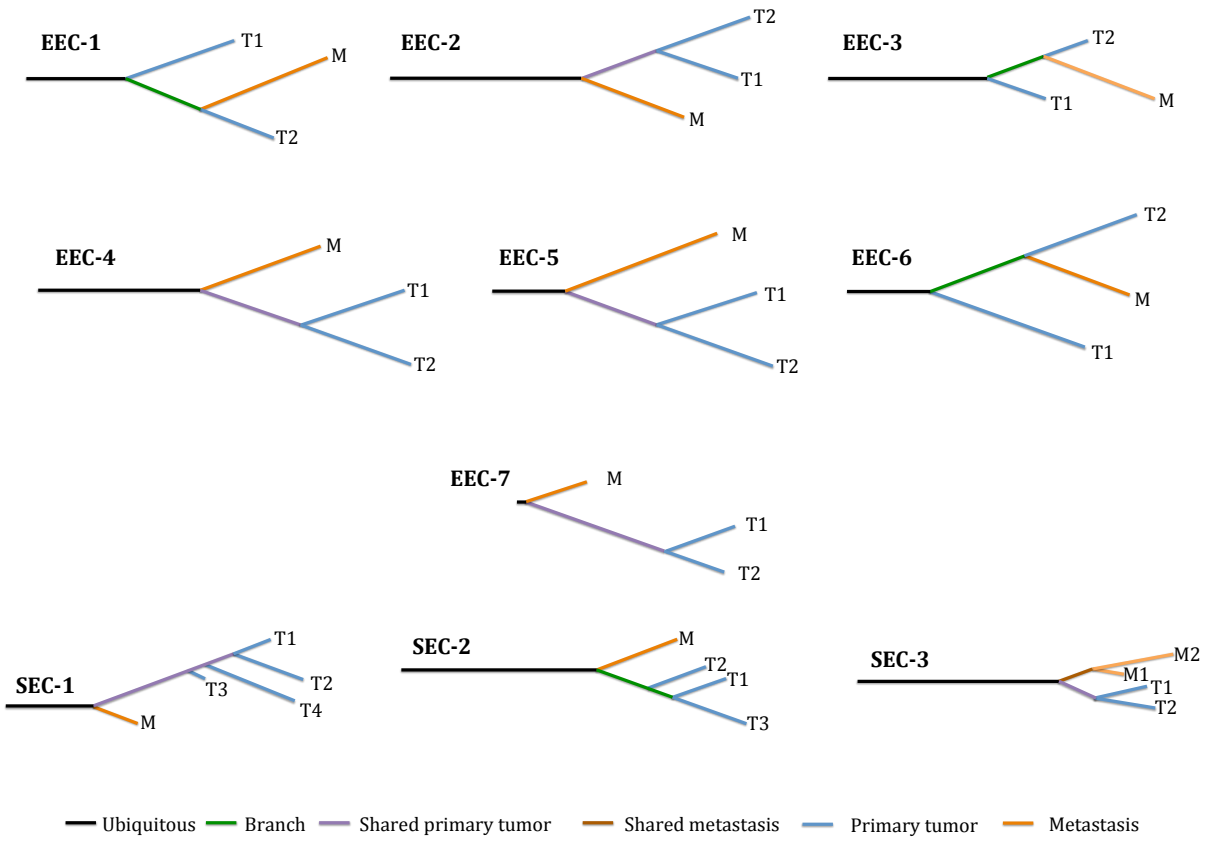


Figure 24 Phylogenetic trees representing tumor evolution in metastatic ECs. Phylogenetic trees were generated using the cancer cell fraction (CCF) mutational data from the WES analysis (see section 3.6.2). Colors represent the samples in which variants were identified: black for ubiquitous variants, green for variants shared between primary tumor and metastatic lesions, purple for variants shared by different primary tumor regions, brown for variants shared between the multiple metastatic regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic samples.

Tumor evolution was analyzed by the construction of phylogenetic trees based on the CCF values for each patient (**Figure 24**). Three different patterns of evolution were observed. The majority of the patients showed a monophyletic pattern, with two main branches representing an initial evolution of the metastatic clone and a subsequent development of additional primary tumor mutations (EEC-2, EEC-4, EEC-5, EEC-7, SEC-1 and SEC-2). SEC-3 patient also showed two

differentiated branches for the primary tumor and metastatic regions, but with a main group of ubiquitous variants that developed in two branches: one of them with shared variants between both primary tumor regions and the other one with the same pattern for the metastatic lesions. On the other hand, in the EEC-1, EEC-3 and EEC-6 patients the metastatic lesion seems to develop from one of the primary tumor regions, indicating the presence of polyphyly. Altogether, these results suggested that different evolution patterns are found in EC progression, apparently not related to histological or molecular classification of the tumor.

4.3.4 Targeted sequencing broadened the view of ITH implication in metastatic ECs

With the aim of validating the genetic variants identified by WES, a subsequent study using targeted massive parallel sequencing was performed in the previously analyzed samples. A total of 596 variants were reanalyzed in the validation process (Supplementary Table 5). Selection criteria included all the pathogenic variants along with additional non-pathogenic variants to reach at least 40% of those detected by WES in the CN-low and serous cases and 10% in the MSI tumors. Seven of the 596 variants (1.2%) could not be analyzed due to a failure of the primers in the amplification step. A validation ratio of 99% was found, being 582 of the 589 analyzed variants detected as previously described by the WES analysis.

In order to expand the study of ITH in these cases, additional FFPE samples from each patient were also subjected to ampliseq sequencing. At least 6 different tumor regions per patient, accounting for multiple primary tumor as well as metastatic locations were included in this analysis (mean = 9, range = 7-11) (**Table 9**). Uterine aspirates (UAs) were also included when available. Targeted sequencing was performed to a mean depth of 1,243x (range = 224-5,431x). No differences were found between the sequencing quality metrics of frozen and FFPE tumor samples (Supplementary Table 17). Targeted sequencing results are included in Supplementary Table 18.

The comparison between those variants identified in the multiple sample analyzed by WES or by targeted sequencing revealed changes in the ITH degree. For this comparison, only those variants included in the validation process were taken into account.

Starting with the CN-low patients (EEC1-4), we observed that the percentage of common SNVs remains similar in two of the cases (EEC-1 and -3) but strongly decreases in the other two (EEC-2 and -4) (**Figure 25a**). In the case of the patients with MSI tumors (EEC5-7), the percentage of shared mutations was higher in the targeted sequencing compared with the exome study (**Figure 25b**). This could be partially explained by the fact that targeted sequencing provided coverage values more than 20-fold higher than WES analysis. In that light, this methodology allows much more reliable DNA sequencing, facilitating the analysis of DNAs as unstable as

those of a tumor with MSI. On the other hand, the three serous cases (SEC1-3) showed a variant distribution similar to that found in the WES analysis, with the majority of them common to all the samples (**Figure 25c**).

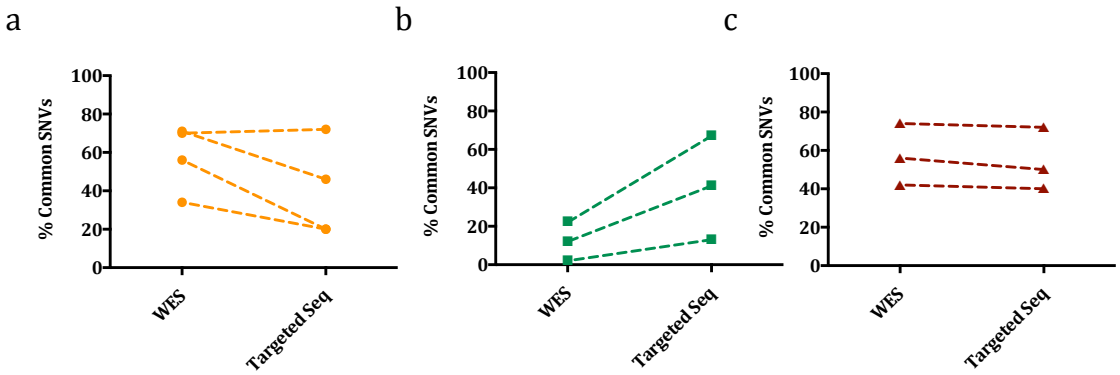


Figure 25 Comparison of the percentage of common variants detected by WES and targeted sequencing validation. Percentage of variants shared by all analyzed samples in each CN-low (a), MSI (b) or serous (c) patients, by WES and targeted massive parallel sequencing analyses. Only those variants considered in the validation process were included in this comparison.

A summary of the targeted sequencing results in each patient is presented in the following paragraphs.

In the EEC-1 case, 12% of the selected variants were common to all samples analyzed by WES while 20% of them were shared by the samples analyzed in the targeted sequencing. This case remained showing a high ITH rate, being genetically similar the T1 and M2 regions and the T4 and M samples (**Figure 26a**). Interestingly, the primary tumor T5 region did not show the majority of the variants detected in the rest of the samples. This could be suggesting that the T5 region is an early clone that contains the initial mutations as those identified in *PTEN*, *CREBBP* or *SIN3A* genes. However, it did not acquire subsequent mutations such as that in *CTNNB1* gene, which was observed in the rest of the regions. These observations are also reflected in the phylogenetic tree obtained for this patient (**Figure 26b**), mainly formed by multiple branches containing between different primary tumor and metastatic regions. Interestingly, the metastatic region M2 appears to be similar to the primary tumor T1 area, both of them being found in an early point of the tumor evolution, while T4 and M samples, after the acquisition of additional variants, appeared at the end of the tree. The oligoclonal clonal composition of this tumor is represented in its polyphyletic evolution.

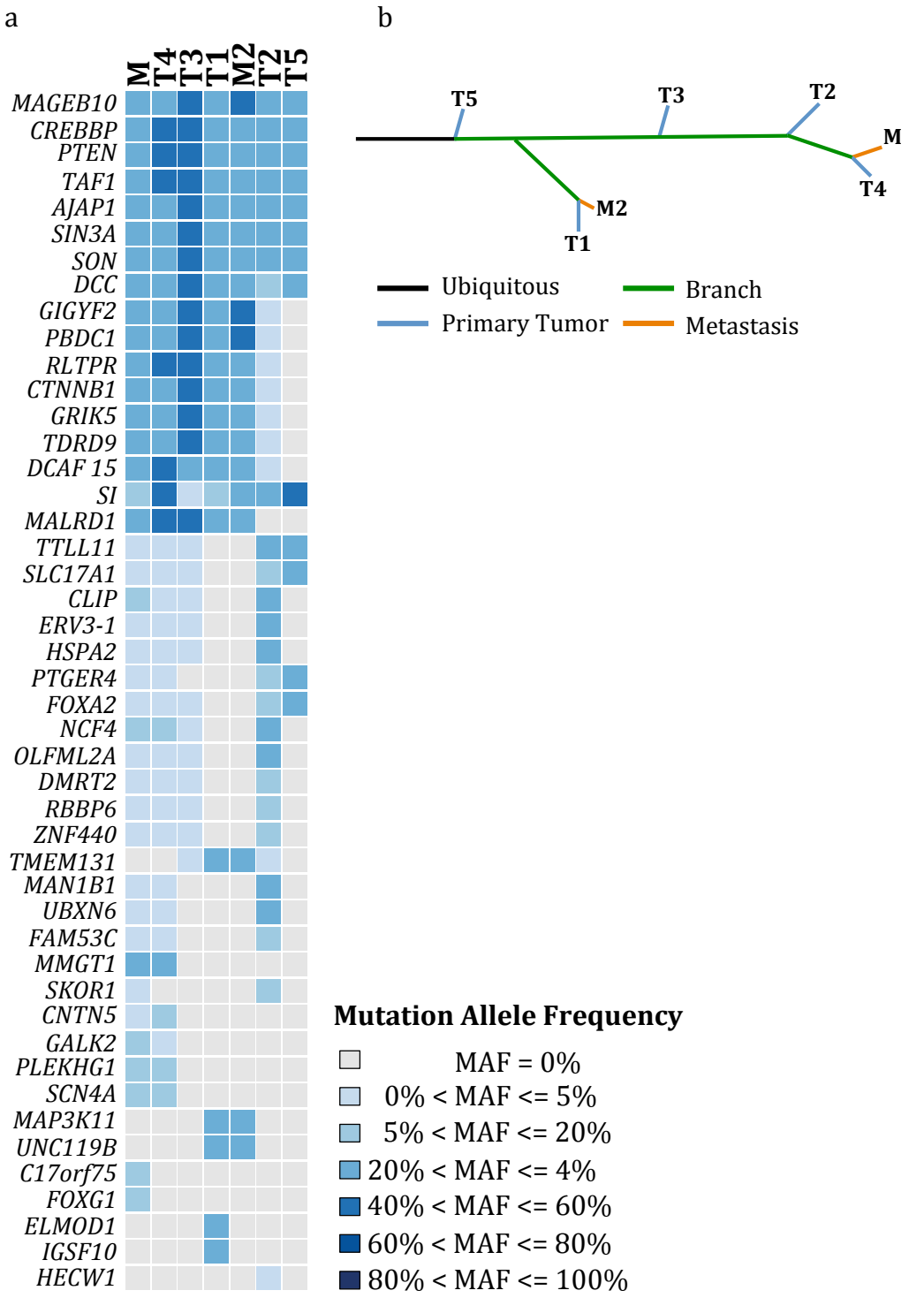


Figure 26 Genetic ITH analysis in the EEC-1 patient. a) Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results. b) Phylogenetic tree generated according to the presence or absence of mutations in the validation analysis. Tumor locations: T1: fundus, T2: left parametrium, T3-T5: different uterine locations from the primary tumor; M: right pelvic node, M2: right primitive node. Metastases were detected and resected at the moment of the tumor diagnosis. Colors represent the samples in which variants were identified: black for ubiquitous variants, green for variants shared between primary tumor and metastatic lesions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic samples.

The EEC-2 case exhibited a lower degree of ITH compared to the EEC-1 patient, although different patterns of variant distribution were also found (**Figure 27a**). While the percentage of shared variants in the WES was 67%, it turned out to be just 20% in the targeted sequencing. In this patient, primary tumor samples T3 and T6 seemed to represent initial clones of the tumor, carrying a low percentage of the variants detected in the rest of the primary tumor regions and being located in an initial point of the tumor evolution (**Figure 27b**). In this case, the *CTNNB1* mutation was found in all the analyzed samples, which could indicate its initiating role in this tumor. No differences were found between the multiple metastatic regions, all localized in the ovary.

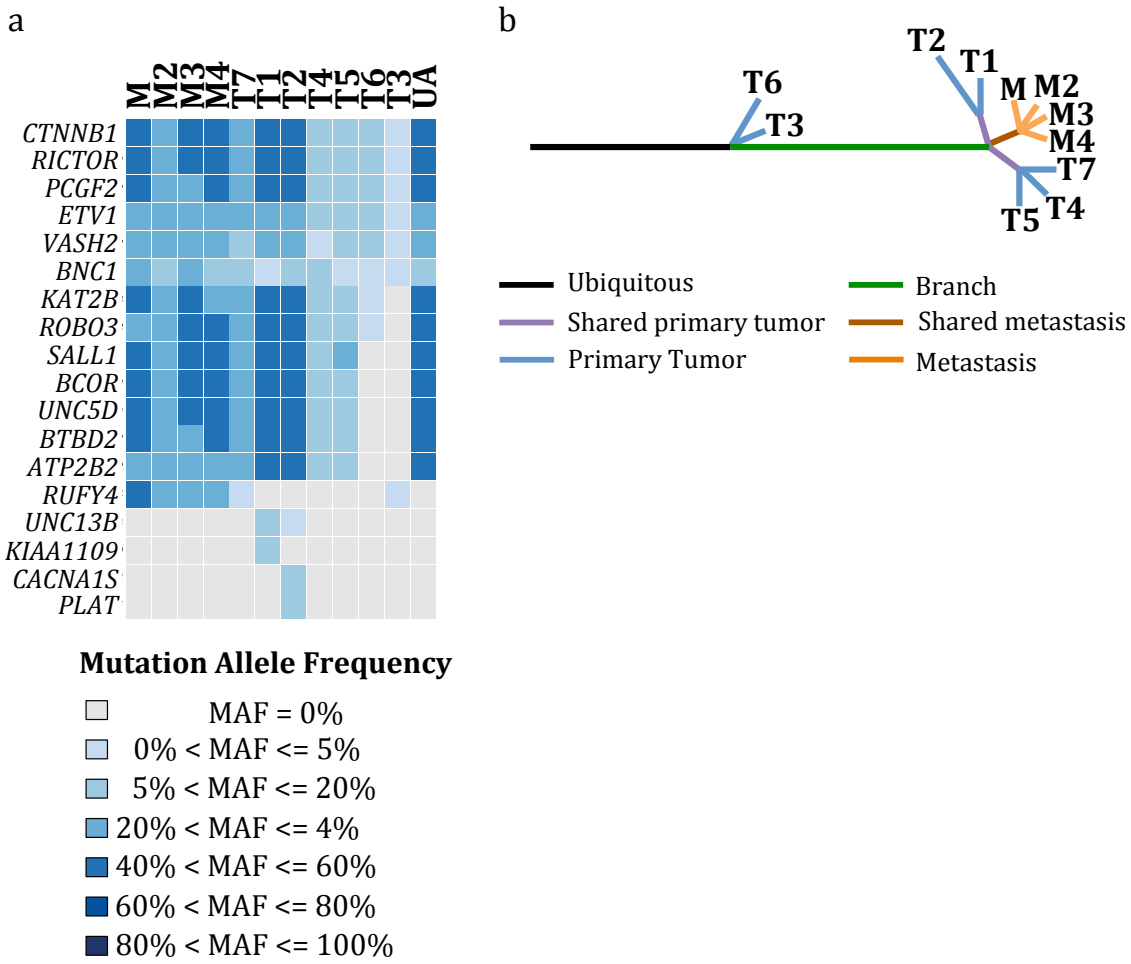


Figure 27 Genetic ITH analysis in the EEC-2 patient. a) Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results. b) Phylogenetic tree generated according to the presence or absence of mutations in the validation analysis. Tumor locations: T1-T7: different uterine locations from the primary tumor; M-M4: different metastatic regions localized in the ovary. Metastases were detected and resected at the moment of the tumor diagnosis. UA: uterine aspirate. Colors represent the samples in which variants were identified: black for ubiquitous variants, green for variants shared between primary tumor and metastatic lesions, purple for variants shared by different primary tumor regions, brown for variants shared between the multiple metastatic regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic samples.

On the other hand, EEC-3 and EEC-4 patients did not show a clear ITH, sharing all the samples the majority of the analyzed variants. Interestingly, in both cases the metastatic regions analyzed belong to a recurrent disease, affecting lymph nodes and the peritoneum in EEC-3 and EEC-4 patients, respectively. In those patients, more than 70% of the variants were common between primary tumor and metastatic samples, while the majority of the remaining variants correspond to additional mutations acquired by the metastatic regions (**Figure 28a** and **b**). This can be observed in their corresponding phylogenetic trees, with a long branch of ubiquitous shared variants and a final expansion of the metastatic regions due to the acquisition of additional mutations (**Figure 28c** and **d**).

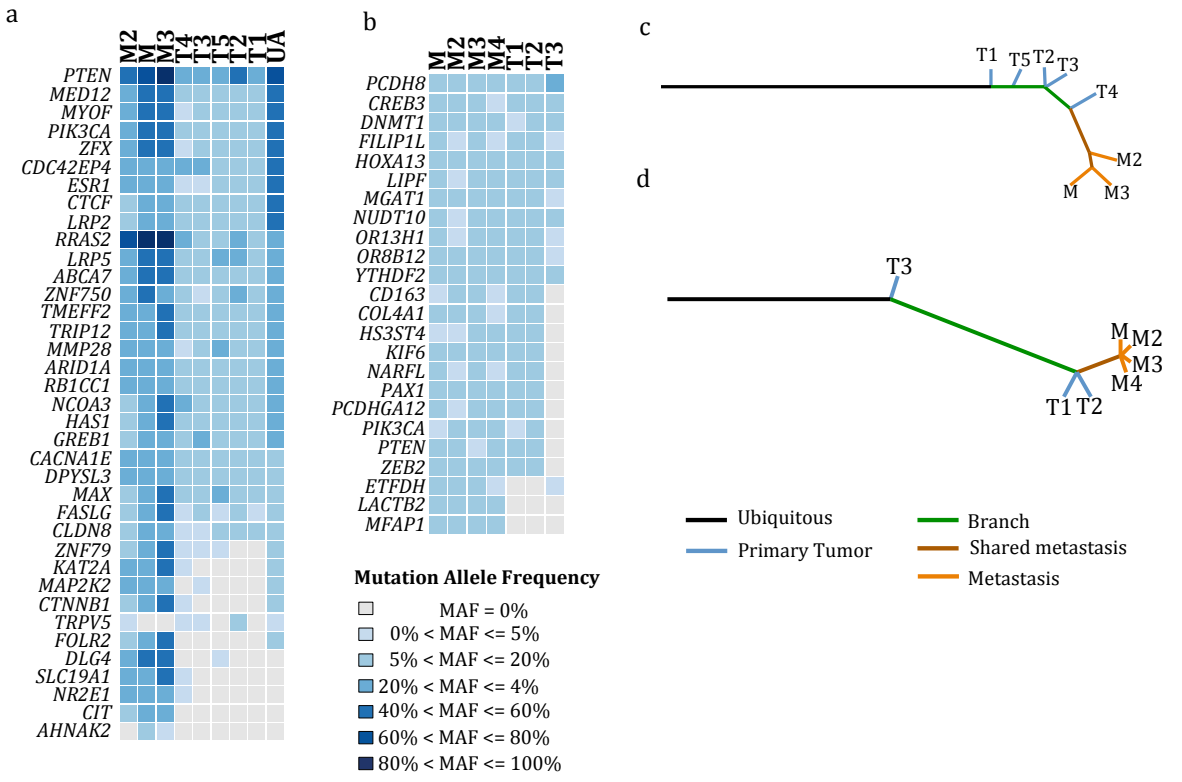


Figure 28 Genetic ITH analysis in the EEC-3 and EEC-4 patients. Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results of the EEC-3 (a) and EEC-4 (b) patients. Phylogenetic trees generated according to the presence or absence of mutations in the validation analysis of the EEC-3 (c) and EEC-4 (d) patients. Tumor locations in the EEC-3 patients: T1-T5: different uterine locations from the primary tumor; M-M3: metastatic regions localized in different lymph nodes. Tumor locations in the EEC-4 patient: T1-T3: different uterine locations from the primary tumor; M-M4: metastatic regions localized in different zones of the peritoneum. UA: uterine aspirate. Metastases were detected and resected 3 years after primary tumor diagnosis in both cases. Colors represent the samples in which variants were identified: black for ubiquitous variants, green for variants shared between primary tumor and metastatic lesions, brown for variants shared between the multiple metastatic regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic samples.

Additionally, UAs that were available from this series of patients were also sequenced (EEC-2_A and EEC-3_A) (**Figure 28**). Interestingly, 72% and 87% of the total analyzed variants were detected in the UAs from EEC-2 and EEC-3 patients respectively. In fact, UA obtained from the EEC-3 patient carried not only the ubiquitous variants but also part of those mostly identified in the metastatic samples. This is the case of the variant detected in *FOLR2* gene that was identified in three metastatic regions (M, M2 and M3) and the uterine aspirate (UA), but not in primary tumor lesions. It is worth mentioning that this aspirate was obtained as a pre-operative biopsy in the first surgery of this patient, being the metastatic lesions resected 3 years later. These results reaffirm the effectiveness of sequencing analysis in UAs for the genetic characterization of EC.

Regarding the patients with MSI tumors (EEC5-7), a higher percentage of common mutations between the multiple samples analyzed in each case was found in the targeted sequencing study when compared to the WES (**Figure 25b**), as previously mentioned.

This is the case of EEC-5 patient, where just 20% of the selected variants used for the validation process were shared by all the samples in the WES analysis, but 41% of them were common by targeted sequencing. In fact, 25 of the 87 (30%) variants determined as exclusive of primary tumor samples by WES, turned out to be also detected in the metastatic region by targeted sequencing. Nevertheless, an important degree of ITH was found, as shown the irregular distribution of the detected variants (**Figure 29a**). The phylogenetic tree of this tumor also show this ITH, with multiple branches and subclones, evolving the metastatic sample from the primary tumor region T3 (**Figure 29b**), indicating the presence of polyphyly.

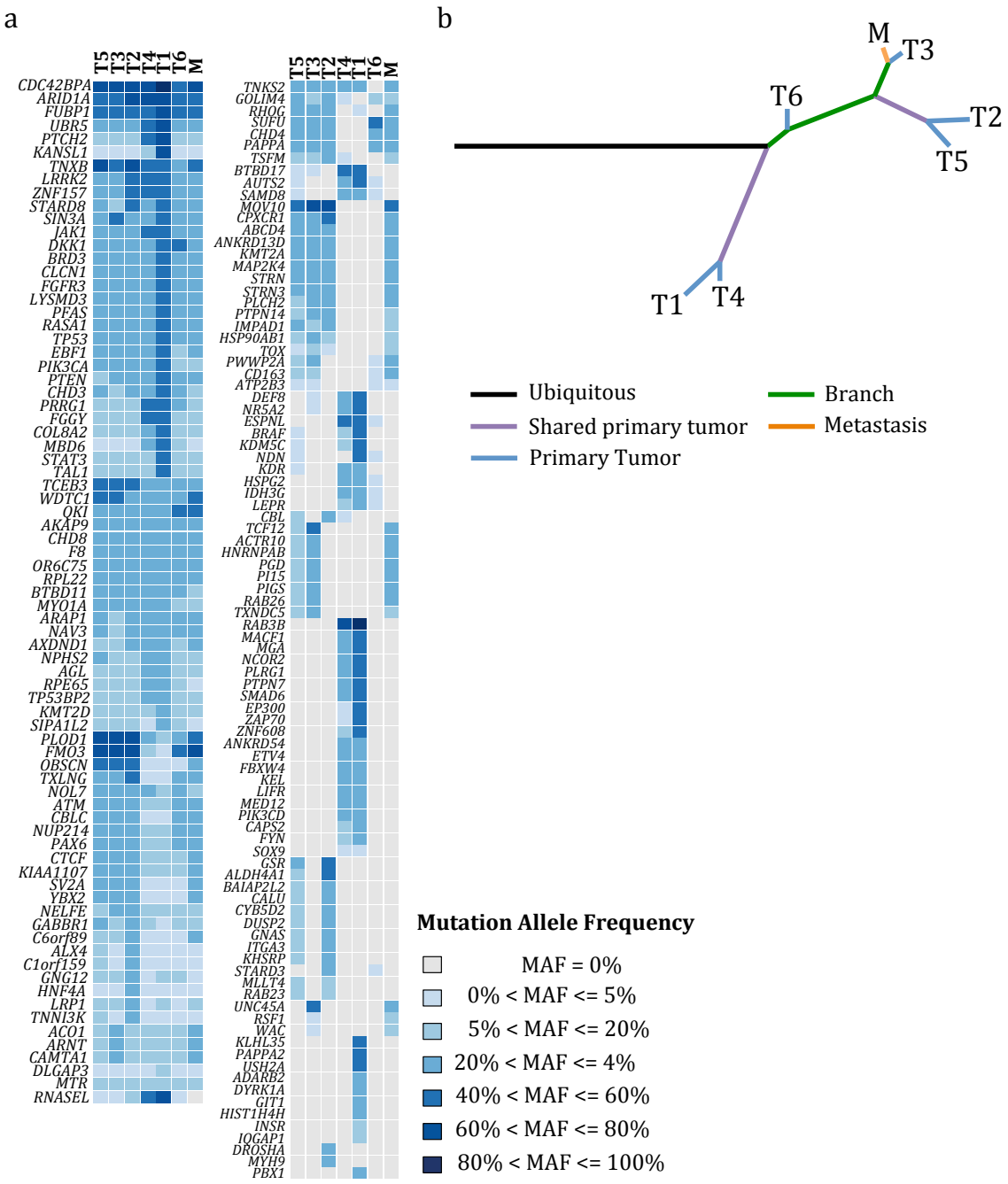


Figure 29 Genetic ITH analysis in the EEC-5 patient. a) Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results. b) Phylogenetic tree generated according to the presence or absence of mutations in the validation analysis. Tumor locations: T1 depth tumor, T2 superficial tumor, T3-T6: different uterine locations from the primary tumor; M: metastatic tumor localized in the peritoneum. Metastasis was detected and resected at the moment of the tumor diagnosis. Colors represent the samples in which variants were identified: black for ubiquitous variants, green for variants shared between primary tumor and metastatic lesions, purple for variants shared by different primary tumor regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic sample.

By contrast, EEC-6 analysis showed a high consistency between the mutational profile of all the analyzed samples, sharing 67% of the analyzed variants (**Figure 30a**) versus 38% found in the WES. In this case, two main branches were found in the phylogenetic tree, distinguishing between primary tumor and metastatic regions in a monophyletic progression of the tumor (**Figure 30b**).

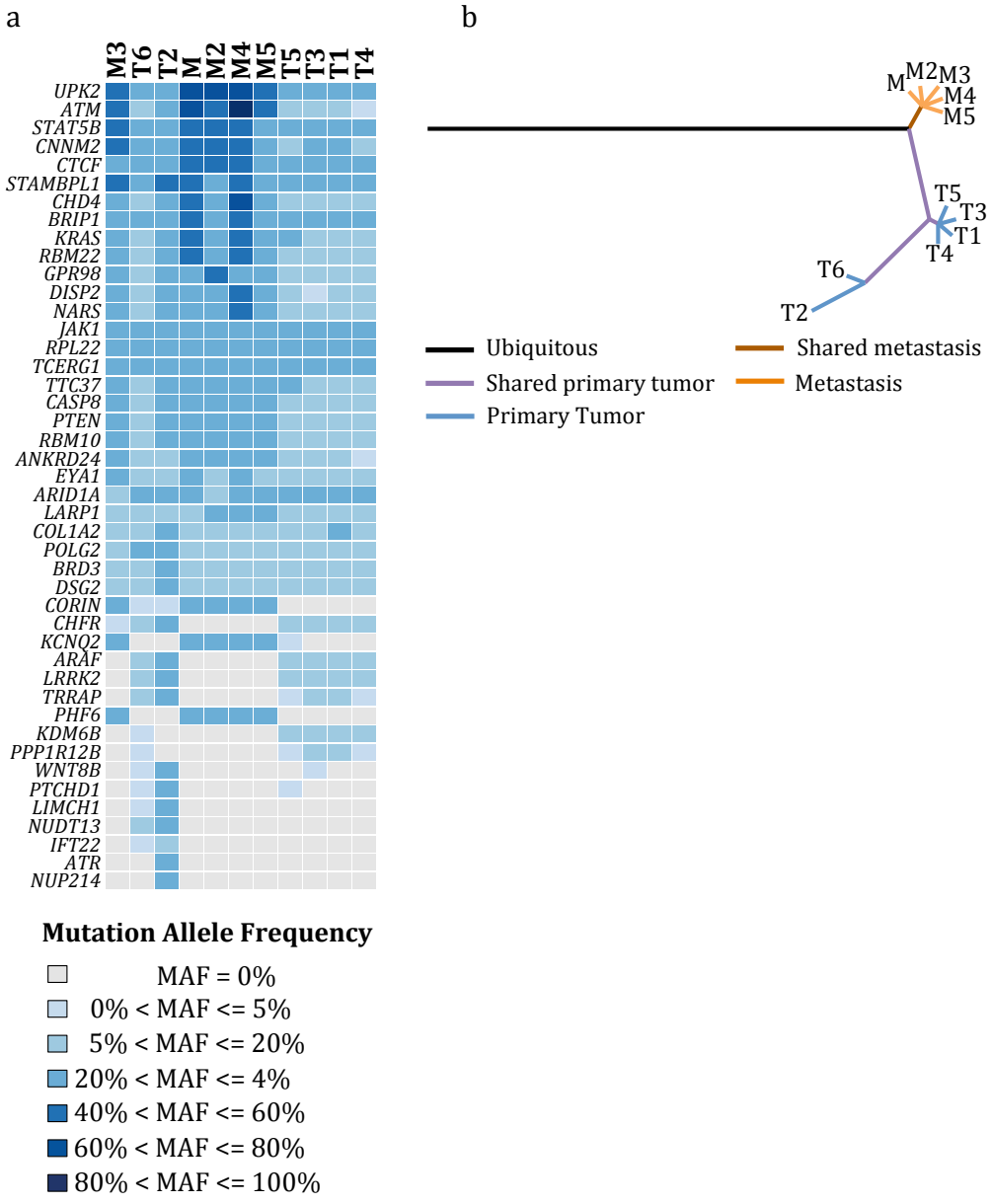


Figure 30 Genetic ITH analysis in the EEC-6 patient. a) Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results. b) Phylogenetic tree generated according to the presence or absence of mutations in the validation analysis. Tumor locations: T1 depth tumor, T2 superficial tumor, T3-T6: different uterine locations from the primary tumor; M-M5: different metastatic region localized in the ovary. Metastases were detected and resected at the moment of the tumor diagnosis. Colors represent the samples in which variants were identified: black for ubiquitous variants, purple for variants shared by different primary tumor regions, brown for variants shared between the multiple metastatic regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic samples.

Nevertheless, although the percentage of common variants in EEC-7 patient was higher with the targeted sequencing than with the WES (13% versus 2%), this value remained very low. Three main patterns of variant distribution were found in this patient: a low percentage of variants shared by all the samples, a few variants mostly identified in the metastatic lesions and additional variants mainly detected in primary tumor regions (**Figure 31a**). The phylogenetic tree generated for this tumor confirmed these assumptions, showing two differentiated branches: the first one nearer to the trunk of the three and composed by the metastatic regions, and the other one with primary tumor regions, which have additional mutations (**Figure 31b**). It is important to note that in this case, the metastatic lesions come from a distant recurrence located at the diaphragm and diagnosed 7 years after the primary tumor was treated. These results could be indicating that during tumor evolution, an early tumor subclone acquired the ability to spread prior to the rest of the primary tumor subclones, which continued gaining mutations not present later in the recurrence samples, suggesting their irrelevant role in the metastatic dissemination of this tumor.

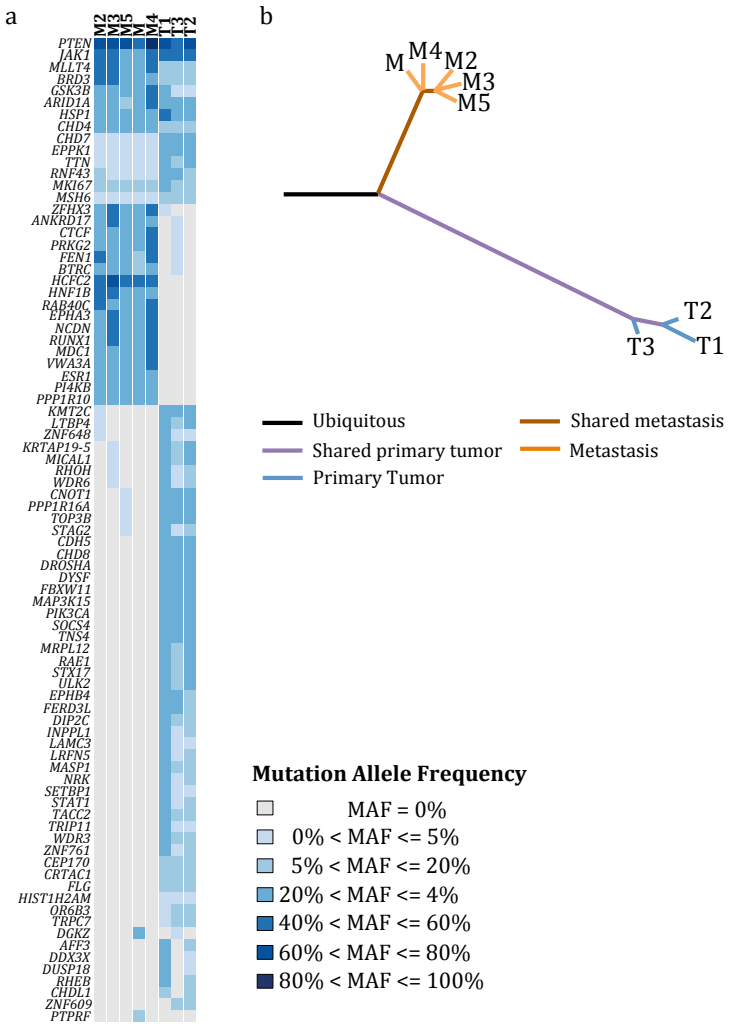


Figure 31 Genetic ITH analysis in the EEC-7 patient. a) Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results. b) Phylogenetic tree generated according to the presence or absence of mutations in the validation analysis. Tumor locations: T1-T3: different uterine locations from the primary tumor; M-M5: different metastatic region localized in the diaphragm. Metastases were detected and resected seven years after primary tumor diagnosis. Colors represent the samples in which variants were identified: black for ubiquitous variants, purple for variants shared by different primary tumor regions, brown for variants shared between the multiple metastatic regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic samples.

In the case of SECs, the majority of genetic variants were consistenly shared among most of the samples. Interestingly, the rest of the detected mutations were mostly identified in the primary tumor regions, indicating an early spread of the metastatic clone to the ovary and a subsequent development of the primary tumor in the uterus (**Figure 32a**, **b** and **c**). Although some mutations were exclusively detected in the metastatic regions, they were a minority. This is the case of the variant identified in *DCP1B* gene in the M1 region of the SEC-2 patient; or the variants in *CSF1* and *TRIM27* genes identified in the metastatic lesions of SEC-3, though in this case minor clones (<1% MAF) were found in some of the primary tumor locations. These results are shown in the phylogenetic trees of each case, in which primary tumor regions and metastatic samples appear separated in two main branches, after a long evolution of the tumor with most of the variants ubiquitously detected (**Figure 32d**, **e** and **f**).

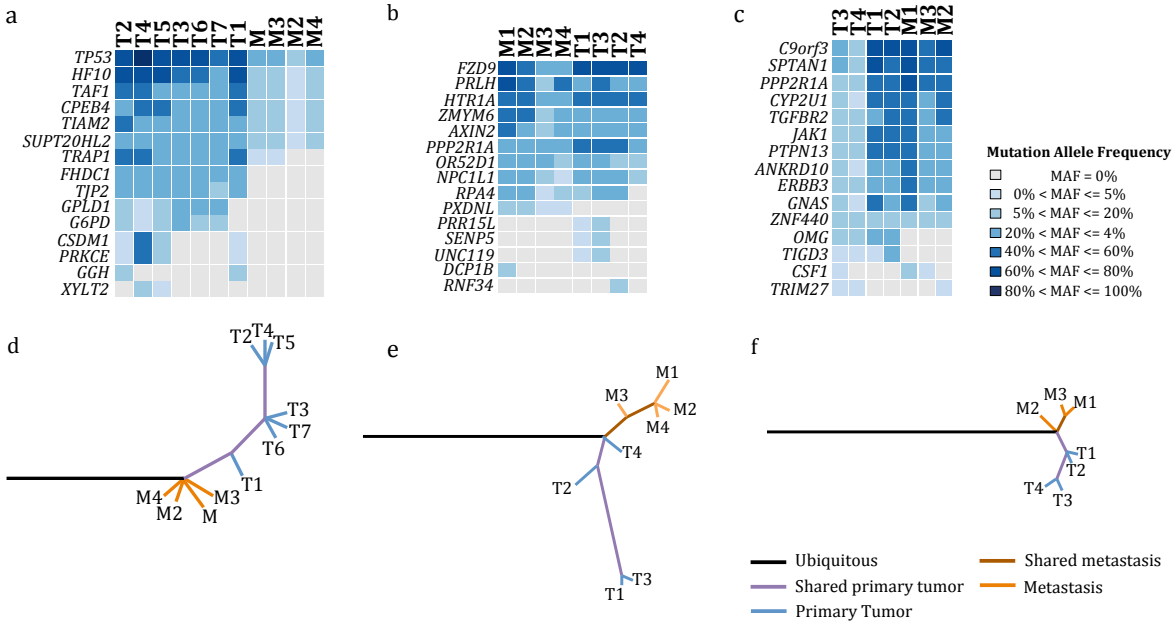


Figure 32 Genetic ITH analysis in the serous-like patients. Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results of the SEC-1 (a), SEC-2 (b) and SEC-3 (c) patients. Phylogenetic trees generated according to the presence or absence of mutations in the validation analysis of the SEC-1 (d), SEC-2 (e) and SEC-3 (f) patients. In the three SEC T regions correspond to primary tumor locations in the uterus, while the metastatic lessions (M) were located at the ovary and ovarian tube. Metastases were detected and resected at the moment of the tumor diagnosis. Colors represent the samples in which variants were identified: black for ubiquitous variants, purple for variants shared by different primary tumor regions, brown for variants shared between the multiple metastatic regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic samples.

In conclusion, different degress of ITH as well as patterns of clonal evolution were found, depending on each analyzed EC patient. It is important to note that the conclusions obtained from the phylogenetic tree studies should be carefully considered since they are based on a targeted sequencing study, not including complete results found in the previous WES analysis.

4.4 Genomic analysis of an ambiguous endometrial carcinoma: the use of new generation platforms and animal models in complex cases

4.4.1 Case description and molecular diagnosis

As previously mentioned, one of the cases included in the WES analysis of metastatic ECs was classified as an ambiguous endometrial carcinoma (AEC). It has been previously described that there is a subset of tumors with discordant microscopic and molecular characteristics that entails real difficulties to be histiologically classified, termed ambiguous tumors[128, 129, 232]. One of the objectives of the present thesis involves the implementation of genetic and genomic studies in order to better characterize and understand this type of tumors.

In this case, the hematoxylin and eosin (H&E) stainning of different regions of the tumor was pointing out towards a carcinoma without readily identifiable features of any of the cells types of differentiated endometrial neoplasms. Sheets of clear, mitotically active tumour cells with areas of necrosis were also observed (**Figure 33a**). Additionally, a previously described immunohistochemical panel of 9 biomarkers used for helping to predict histological subtype[233] was analyzed in this tumor (**Figure 33b**). According to the immunophenotype of this tumor, more than 5 conditions satisfied to predict SEC versus differentiated EEC (grade 1-2). Nevertheless, the tumor showed heterogeneity and in other sections less than 5 conditions were satisfied, not being able to discriminate between SEC or EEC. Therefore, the histological diagnosis was of an AEC.

Additionally and to further characterize this tumor, patient derived xenografts (PDXs) were generated from this AEC (see **3.11** section). For this approach, different tumor regions from the primary surgery were implanted in five different mice, including areas from the superficial (AEC_PDX1) and deep (AEC_PDX2) tumor, right (AEC_PDX3) and left (AEC_PDX4) lymphatic nodes and cervix implants (AEC_PDX5). After tumor growth, samples were resected and analyzed by H&E staining, showing similar histology that the previously found in the patient (**Figure 33c**).

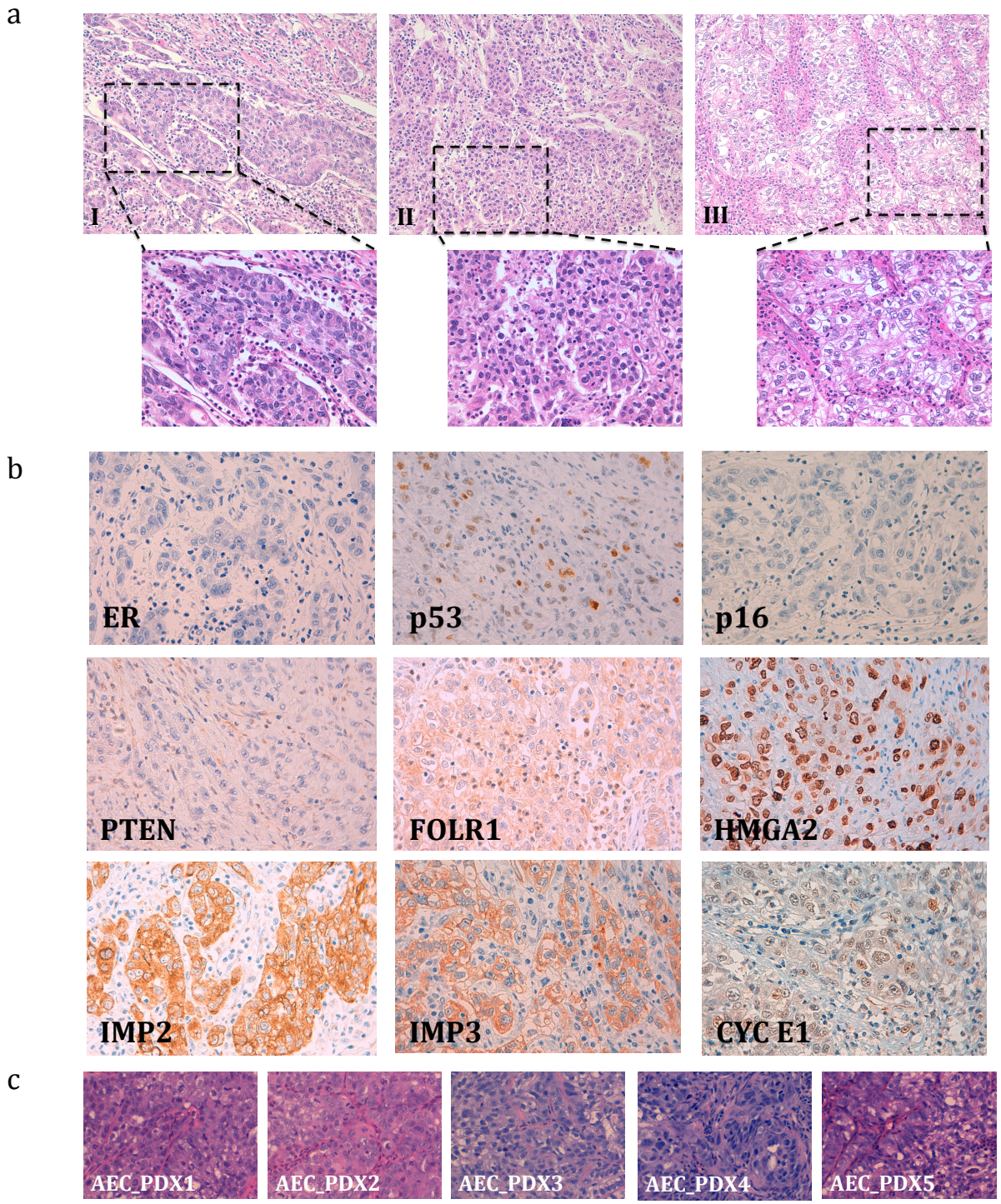


Figure 33 Histological and immunohistochemistry profile of an ambiguous endometrial carcinoma (AEC). a) Representative H&E images of different regions from the primary tumor (I-II) and metastatic lesions (III) of the AEC. Magnification at 20x in the upper panel and histological details of the images in the lower panel (40x). b) Immunophenotype of the AEC based on a 9-protein (ER, p53, p16, PTEN, FOLR1, HMGA2, IMP2, IMP3 and CYC E1-clycin E1) biomarker signature previously described[233] to facilitate the discrimination between SEC and differentiated EEC. Magnification 40x. c) Representative H&E images of the multiple patient derived xenografts (PDX) samples obtained from superficial tumor (AEC_PDX1), deep tumor (AEC_PDX2), right lymphatic node (AEC_PDX3), left lymphatic node (AEC_PDX4) and cervix implants (AEC_PDX5) of this patient. Magnification 40x.

In order to characterize this tumor at a molecular level and for better understanding its clinical behavior, we studied it by WES to identify somatic mutations and copy number alterations. Regarding its controversial origin, we analyzed its genetic profile using the TumorTracer server (see *3.5.4.1* section)[219], which compare our tumor genomic features with a dataset of more than 200 tumors. These results indicated that endometrium was the tissue of origin in this tumor, showing the highest classification score (0.548) (**Figure 34a**). A confidence score of 0.246 was obtained for the endometrial cancer diagnosis. For this score, a predictive positive value of 79% was obtained when comparing to the TumorTracer dataset, with 11 of the 14 positive cases properly classified (**Figure 34b**). These results supported the pathological diagnosis of this tumor.

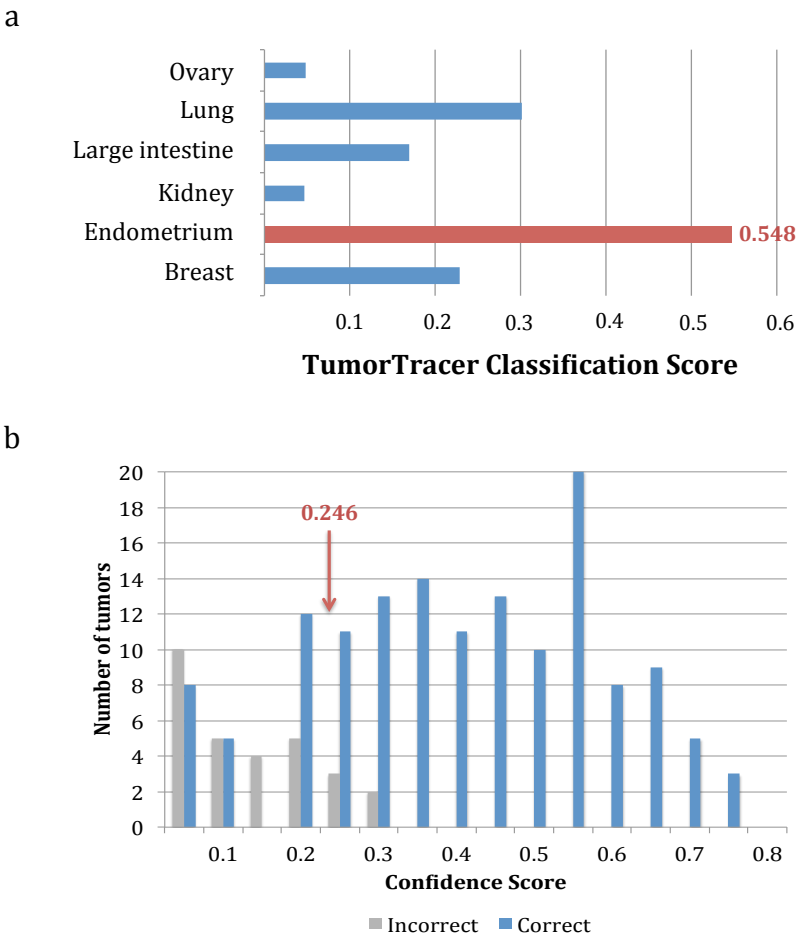


Figure 34 TumorTracer results based on WES analysis of the ambiguous endometrial carcinoma (AEC). a) TumorTracer classification score identified the endometrium as the most probable tissue of origin. b) Representations of the number of tumors from the TumorTracer datased incorrectly (grey) or correctly (blue) classified depending on the confidence score (see *3.5.4.1* section for further information). TumorTracer classification and confidence scores obtained for our patient are marked with a red arrow.

4.4.2 WES analysis revealed uncommon results in the molecular characterization of the ambiguous endometrial carcinoma

Although this case was molecularly classified as an endometrial carcinoma by the TumorTracer server, WES results revealed it did not show any of the most common mutated genes in EC. In fact, the mutational signature 13, which is characterized by C>G and C>T changes and has been attributed to the AID/APOBEC family of cytidine deaminases[217], was identified in the three samples analyzed by WES (T1, T2 and M) (Supplementary Table 16). This mutational signature has been previouly detected in bladder and breast cancer[16], but to our knowledge it has not been described in endometrial carcinomas until now.

Genetic and genomic variants distribution between analyzed samples of the AEC were studied. As previously observed in other cases, the percentage of common SNVs was considerably higher compared to the shared genes altered by SCNAs (81% versus 10%) (**Figure 35a** and **b**). In contrast, 37% of the genes affected by SCNAs were common to primary tumor samples (T1 and T2). This similarity between the primary tumor regions was reflected in the phylogenetic tree obtained for this patient, which showed two main branches, one of them including the primary tumor samples an the other one the metastatic sample (**Figure 35c**), indicating a monophyletic evolution programme. The low presence of ITH at the genetic level was also revealed by the CCF study, that identified 80% of variants in a clonal status. CCF heatmap (**Figure 35d**) also showed three main subgroups of variants distribution: the main subgroup was composed by variants common to all samples, the second one comprised those variants shared by the primary tumor samples (T1 and T2), while the rest of variants were identified in a single sample (T1, T2 or M). Moreover, only 24 of the variants identified by WES were mutations with a potentially pathogenic consequence in the corresponding protein, and only one (in the *ATM* gene) with a pathogenic consequence (Supplementary Table 14, **Figure 35d**). As previously mentioned, none of the genes carrying these variants are considered frequently mutated in EC. Nonetheless, a copy-number gain of the gene *CCND1* was found in all samples. *CCND1* gene codifies for the cyclin D protein, which have been described to be amplified in 26% of ECs[234, 235].

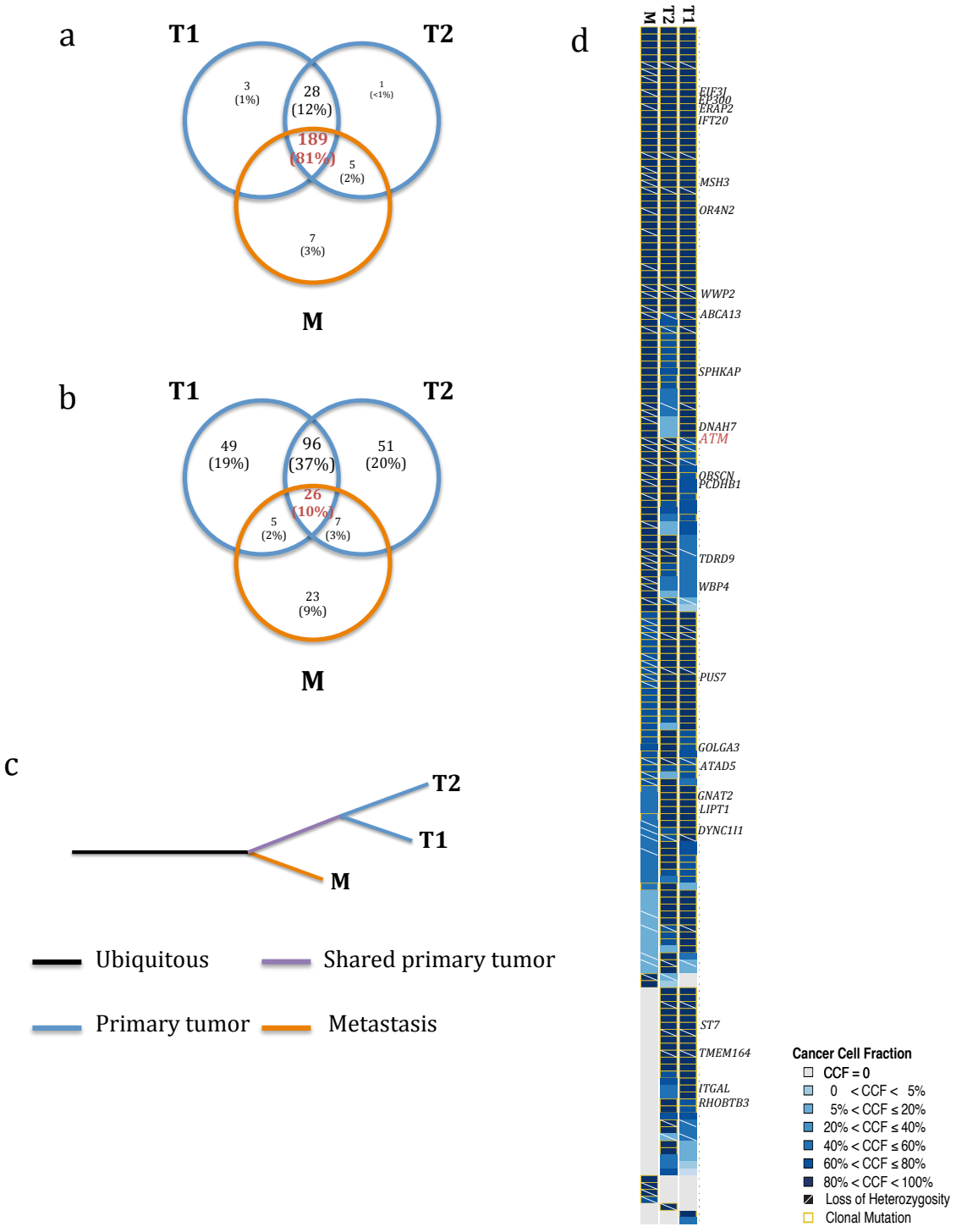


Figure 35 WES results obtained in the analysis of the ambiguous endometrial carcinoma. a) Venn diagram representing the genetic variant distribution found in the WES analysis of an ambiguous endometrial carcinoma (AEC). b) Venn diagram showing the somatic copy number alteration distribution found in the WES. c) Phylogenetic tree symbolizes tumor evolution based on the WES results (see 3.6.2 section). d) Heatmap indicating the presence (blue) or absence (grey) of the mutations identified by WES. Different blue colors indicate the cancer cell fraction (CCF) value obtained for each variant (see *3.5.4.3*). Loss of heterozigosity (LOH) is represented by a diagonal white line in the corresponding variant box. The boxes of those variants identified as clonal are labeled with a yellow square. Genes with a probably pathogenic consequence are indicated in the heatmap. *ATM* gene, marked in red, was the only one with a totally pathogenic mutation.

4.4.3 WES results validation in additional tumor samples and genetic analysis of patient derived xenografts (PDX) models

To further characterize the ITH in the AEC, a selection of 86 of the variants identified in the WES were analyzed in a validation process using targeted ampliseq sequencing (Supplementary Table 5). These variants were selected in order to include all the potentially pathogenic and additional non-pathogenic variants to cover 40% of those detected by WES. In addition to those samples previously analyzed by WES, three additional FFPE primary tumor samples as well as the uterine aspirate tissue were included in the analysis (**Table 9**).

Bioinformatics analysis revealed a higher degree of genetic ITH in the targeted sequencing analysis than that previously observed in the WES. Only 17% of the variants were shared between all the samples in the validation process, against 69% observed in the selected variants from WES. This is mainly due to the fact that primary tumor region T4 only showed a small percetange of the variants, probably indicating that it is an initial clone of the tumor. On the other hand, primary tumor regions T1 and T2 and the metastatic lesion (M) appeared at the end of the three, showing the final part of the tumor evolution (**Figure 36a** and **b**). Besides, a 81% of the analyzed variants were detected in the UA sample, confirming its utility capturing ITH of endometrial tumors (**Figure 36a**).

Additionally, five tumor samples from patient derived xenografts (PDX) were included in the targeted sequencing process in order to investigate its reliability representing the genomic landscape of the patient's tumor (**Figure 33c**). A total of 79 of the 86 analyzed variants (81%) were also found in the five PDX samples (**Figure 36a**). Only 11 variants were not detected in any PDX sample, while the other 5 variants were present in some of them. These partially or non-detected variants were previously found in the patient primary tumor regions with a heterogeneous distribution. Therefore, these results suggested that PDX models could be not representing the genetic ITH found in the different tumor samples of this patient.

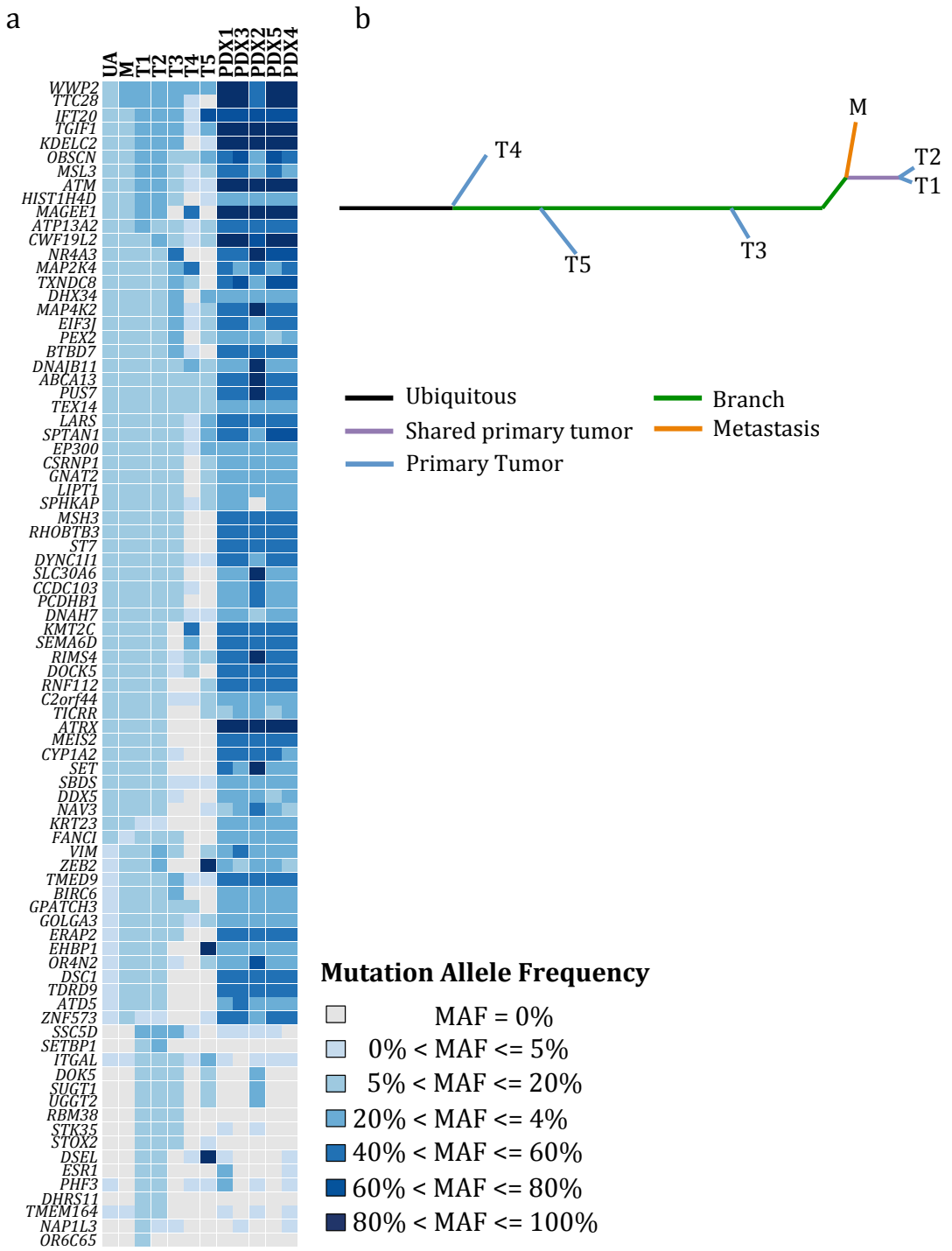


Figure 36 Targeted sequencing results obtained in the multiple samples analysis of the AEC patient and PDX models. a) Mutation Allele Frequency (MAF) heatmap based on the targeted sequencing validation results. b) Phylogenetic tree generated according to the presence or absence of mutations in the validation analysis. Tumor locations: T1-T5: different uterine locations from the primary tumor; M: lymphatic node. Metastasis was detected and resected seven years after primary tumor diagnosis. UA: uterine aspirate. Colors represent the samples in which variants were identified: black for ubiquitous variants, green for variants identified in primary tumor and metastasis regions, purple for variants shared by different primary tumor regions, blue for variants exclusively observed in primary tumor samples and orange for variants only detected in the metastatic sample.

4.4.4 WES analysis in patient derived xenografs (PDX) and comparison to the patient mutational profile

Since targeted sequencing used in the validation of the patient WES study showed genetic homogeneity between the different PDXs, WES analysis was also perfomed in those PDXs and in a normal mouse tissue used as control (see **3.11** section).

Bioinformatics analysis of PDX WES identified a total of 393 variants (Supplementary Table 19). Interestingly, only 197 of the 393 (50%) variants had been previously observed in the patient samples, being most of them common to all the PDX samples (187/197, 95%). (**Figure 37**). A total of 36 variants of the 233 (15%) previously detected in the WES of the patient were not identified in the PDX samples, the majority of them non-common between patient samples (25/36, 69%).

From the new variants detected in the PDX samples, only 41% of them were shared by all the samples (**Figure 37**). These results confirmed that PDX samples carried the majority of variants previously identified in the patient, although not those with a heterogeneous distribution. Moreover, the number of variants considerably increased with the tumor growth during the PDX generation. Consequently, the genetic profile of the PDX samples significantly differs from the original one found in the patient.

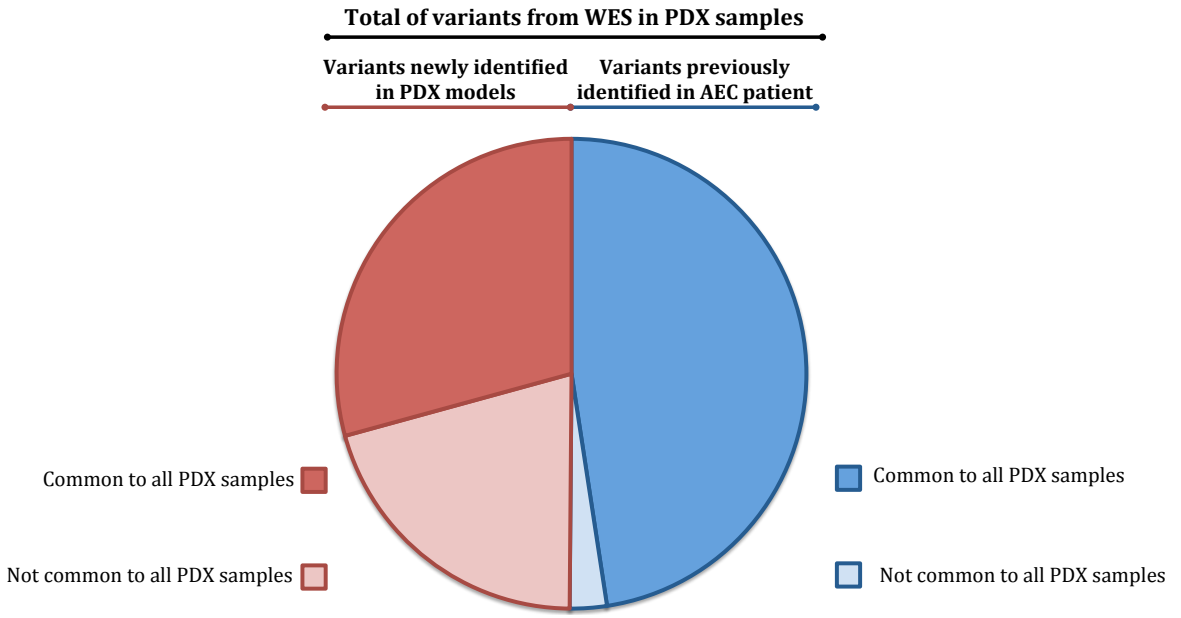


Figure 37 WES variants identified in the patient derived xenografts (PDXs) obtained from an ambiguous endometrial cancer (AEC) patient. Circle chart showing the percentage of variants detected in the WES of PDX samples. Red colors represent those variants newly identified in the PDX models, while blue colors represent those previously detected in the patient study. Dark red or dark blue colors marked those variants shared by the five PDX analyzed samples.

4.4.5 Personalized treatment based on WES data obtained from an ambiguous endometrial carcinoma: validation in a preclinical model

Finally, we analyzed the mutational profile found in the AEC in a pharmacological context in order to evalutate the potential relationship between the mutational status and available drugs. For this purpose we performed an *in silico* study using different drug databases (CTD (http://ctdbase.org/) and STITCH (http://stitch.embl.de/)[220-222]). This analysis (*3.5.4.6* section) revealed different treatments which could have a potential effect regarding the molecular alterations identified in this case, reporting bortezomib and paclitaxel as the most appropiated (**Figure 38a**). Small pieces of the AEC_PDX1 model were surgically transplanted subcutaneously into a second generation of mice. PDX models were treated following the *in silico* study results with bortezomib as a single agent (0.25mg/kg) and a combination of bortezomib plus paclitaxel, in comparison to the standard treatment used in this type of cancer (15mg/Kg paclitaxel –PTX- and 50 mg/kg carboplatin –CBTP-) and placebo. For testing effectiveness of these treatments two mice were used in each condition as indicated in **3.11** section. Mice were treated during 30 days with specified drugs and tumor growth were monitored twice weekly during 70 days (**Figure 38b**). Although these are preliminary data, we observed a statistical tendency (p=0.11) towards decreased tumor growth in the bortezomib-PTX when compared to standard PTX-CBPT treatment, suggesting the applicability of this treatment in this particular clinical context as well as the potential utility of this *in silico* study based on WES data.

a

Mutated genes	Related drug	Data base
CYP1A2, TICRR, MAP2K4, GSPT1, ATM, DYNC1I1	Bortezomib	CTD
CYP1A2, MAP2K4	Bortezomib	STITCH
CYP1A2, MAP2K4, NAV3, TOP2B	Paclitaxel	CTD
CYP1A2, TICRR, MAP2K4, GSPT1, ATM, DYNC1I1, NAV3, TOP2B	Bortezomib Paclitaxel	CTD/STITCH

b

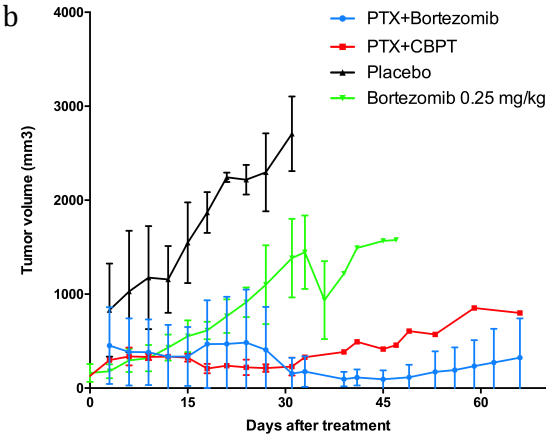
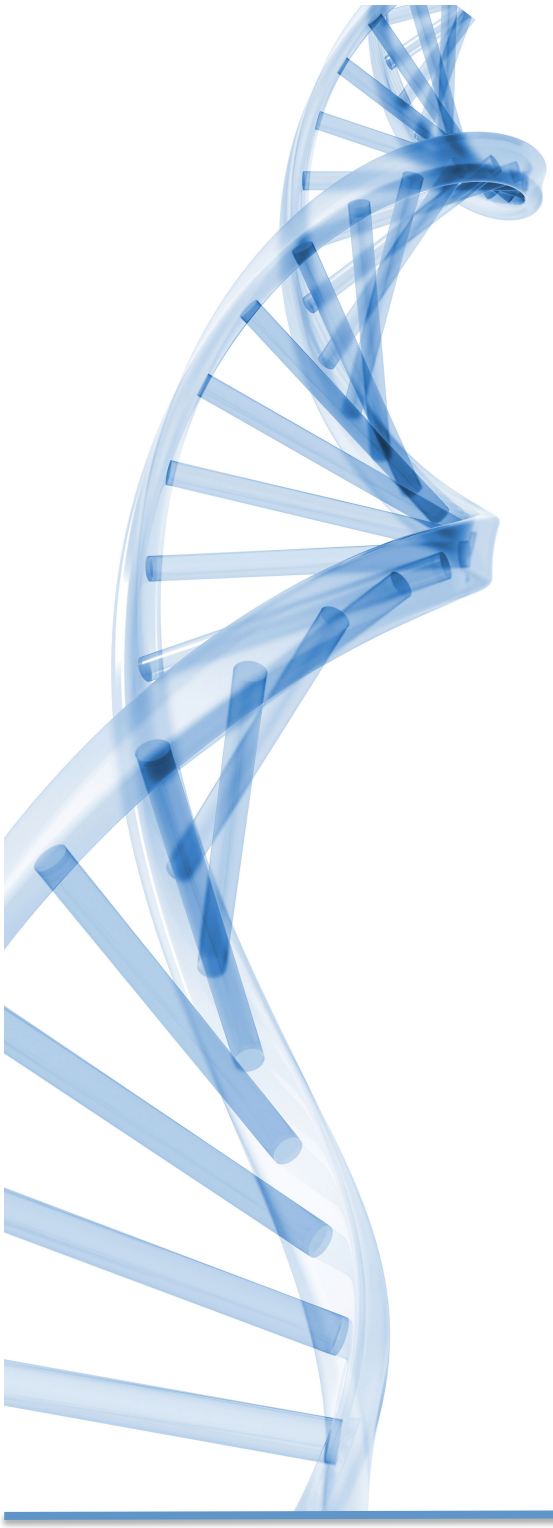


Figure 38 *In silico* study of drug treatment based on WES results and its application in PDX models. a) Selection of treatment regarding the *in silico* study based on information available in CTD (http://ctdbase.org/) and STITCH databases (http://stitch.embl.de/) and mutated genes found in AEC. b) Tumor growth representation of PDXs obtained from AEC treated with indicated drug conditions during 30 days. Tumor volumes were measured twice weekly. (Mann-Whitney test, p=0.11). Data show the mean of 2 independent mice for each treatment; bars, s.d. PTX: paclitaxel, CBPT: carboplatin.



DISCUSSION

5 DISCUSSION

The presence of intratumor heterogeneity (ITH) at a phenotypic level has been described since the early days of cancer research, leading to a wide investigation about the clonal origin of human neoplasms as well as the role of molecular heterogeneity in cancer progression[20-22, 236]. The development of high-throughput techniques, and especially next-generation sequencing (NGS), made a real difference in the knowledge about ITH[237-239]. Thereby, different methodological approaches have provided deeper insights into ITH. The use of bioinformatics tools to infer the subclonal population of an individual biopsy of tumor[215, 240, 241] was followed by the implementation of multiregion analyses[23-25] and more recently of single-cell sequencing[242, 243].

While the majority of studies revealed a branched evolution within tumor, supporting its complex structure based on multiple genetically diverse subclones, linear patterns of tumor evolution have been also found[42, 46]. In these cases, the acquisition of advantageous mutations give rise to new clones that replace the previous ones, producing a relative homogeneity when a single biopsy is analyzed.

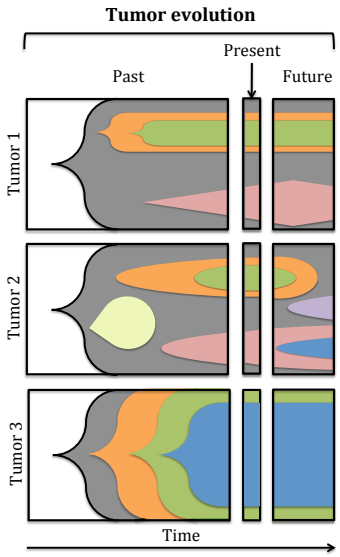


Figure 39 Examples of tumor evolution types. Three different examples of evolution patterns are represented in the graphic. Left panels show the tumor development prior to the sample analysis (past). Middle panels represent the specific moment when the sample is obtained for sequencing analysis (present). Right panels show the possible evolution of each tumor (future). Tumors 1 and 2 show a branched pattern, while tumor 3 displays a linear evolution. Although tumor 1 and 2 may have the same present heterogeneity distribution, their past and future progression have different patterns. Colored shapes represent subclonal populations of the tumors. Modified from Hiley *et al*[239].

Different examples of tumor evolution are schematized in **Figure 30**. Regarding these patterns, while tumor 1 and 2 have a branched distribution, tumor 3 shows a lineal progression along the time. However, if we only analyze the current status of each clinical case with a single biopsy throughout the development of the tumor, sequencing analysis may suggest a similar pattern and clonal composition in tumors 1 and 2, although they have a different history of evolution[239]. These differences found in tumor evolution

highlight the importance of analyzing the spatial as well as temporal growth of the tumors and their dynamic acquisition of alterations.

Numerous studies have shed light on the extent of tumor diversity, including several solid tumor types such as pancreatic[24, 26], renal[25, 27], lung[28-30], breast[31-33], colorectal[34, 35], glial[36-38] and prostate[39, 40], among others; as well as hematologic malignancies[41-46].

Similar analyses have been also reported about gynecological cancers[47], most of them focused on ovarian carcinomas, which have the highest mortality rates among this kind of tumors[58]. However, there are other tumor subtypes less explored, as it is the case of endometrial carcinoma (EC)[50], which is the most common gynecological malignancy in the developed countries, but have a better prognosis compared to ovarian carcinomas[58, 120]. Interestingly, several studies have analyzed ITH in tumor progression in other tumor types, exposing heterogeneous patterns in driver and targetable mutations between primary tumor and metastatic or recurrent disease[24, 25, 28, 38, 115]. In this regard, it has been demonstrated that oncologic treatment could act as a selection pressure in some malignancies[42, 244, 245]. Nevertheless, genomic instability is considered the major cause of cancer heterogeneity nowadays, leading to an increased mutation rate and generating distinct genomic footprints known as mutational signatures[16, 246]. Although genomic instability is present in most solid tumors and hematopoietic malignancies, multiple pathways have been implicated in its development, including defects in DNA repair mechanisms, DNA replication or chromosome segregation among others[246, 247].

To characterize the occurrence of ITH in different gynecological cancers and to assess its implication in metastatic and recurrent disease, in this thesis we have focused in the following specific **objectives**: a) ITH analysis in a *TP53*-null ovarian carcinoma patient, b) endometrial ITH characterization, c) application of exome studies for deeping on an ambiguous endometrial carcinoma and d) implementation of genetic analysis in uterine aspirates for capturing ITH existent in endometrial tumors.

The presence of ITH has been extensively described in **ovarian cancers**, especially in high grade serous ovarian carcinomas (HGSOCs)[110-118], where *TP53* mutations are broadly present[89, 90]. Mutation in *TP53* is considered the main cause of the widespread genetic instability found in HGSOCs[61]. Despite the majority of them are missense mutations that lead to protein accumulation; around 30% of somatic variants are considered null mutations, including nonsense, frameshift or splicing junction variants that cause the complete absence of the protein[94]. In this sense, while it has been described that not all *TP53* missense mutations have equivalent consequences in the protein function[248], little is

known about null mutations. Furthermore, during the last years the potential impact of *TP53* mutational status as a prognosis factor have been widely analyzed, though conflicting results make it still a controversial issue[103].

On the basis of these precedents, we first performed a meta-analysis using the genomic data generated by the TCGA consortium[90], classifying patients regarding their *TP53* mutational status in mutated, null or wild-type tumors. This analysis revealed significant differences in the fraction of copy number alterations between the *TP53*-null and wild-type tumors, which could be suggesting the existence of diverse mechanisms underlying genomic instability or at least a different level of effect (**Figure 9a**). However, although from the genomic instability point of view *TP53* altered ovarian cases seem to be similar, *TP53*-null subgroup showed an intermediate overall and progression free survival between mutated and wild-type tumors (**Figure 9b**).

Consequently, we aimed to **analyze ITH in the *TP53*-null tumor subgroup**, less characterized than the *TP53*-mutated tumors. For this purpose, whole-exome sequencing (WES) and comparative genomic hybridization (CGH) analyses were performed on multiple regions from a patient diagnosed with a *TP53*-null recurrent HGSOC. As previously reported in other ovarian cancer studies[113-117], significant heterogeneity was found not only at the genetic level but also at the somatic copy number alterations (SCNAs) (**Figure 11a** and **12a**). This divergence was consistent among the analyzed samples, being the recurrence regions as similar between them as with the primary tumor. Mutational analysis by Sanger sequencing of the selected variants in additional tumor samples from the same patient revealed the clonal composition and phylogenetic evolution of this carcinoma (**Figure 13a** and **b**). These results were consistent with a situation in which the primary tumor is composed by mutational heterogeneous clones, with some of them giving rise to the recurrence lesions. In this case, ITH seemed to be intrinsic to the primary tumor and not a consequence of the therapy, which is consistent with previously published data[113]. Hierarchical clustering of tumor regions based on the presence/absence of mutations allowed to dissect the tumor clonal evolution (**Figure 13b**). This analysis showed that different recurrence regions were closely related to distinct and specific primary tumor subclones, which suggests an polyphyletic evolution as it has been previously described in other tumor types[24, 40]. However, these conclusions could be biased because this study was performed using preselected variants. Next, to assess the biological pathways that could be altered in this tumor, functional annotation and network analysis were carried out using the mutational profiles (**Figure 14**). Functions intrinsic to tumor growth and spread, such as cell proliferation, cell cycle control and cell adhesion appeared consistently mutated, while other less expected

pathways, such as microtubule-based movement and lipid metabolism could also be affected. Microtubule function is essential for chromosomal partitioning and segregation[249], thus, alterations in microtubule cytoskeleton might be in part responsible for the wide genomic and chromosomal aberrations that HGSOCs usually show[90]. Lipid metabolism is frequently altered in cancer disease[250], and interestingly, some of the specific detected mutations have been previously involved in tumor progression in other tumor types[251-255]. In this sense, our study revealed that these gene alterations were mainly shared between the three samples analyzed by WES (P1, ER1 and IR1), suggesting a founder and clonal role for lipid metabolism deregulation in this HGSOC. Furthermore, since *TP53* has been shown to be a key regulator of lipid metabolism[256], altogether, these results suggested that lipid metabolism could be considered as a therapeutic target in this specific *TP53*-null patient, although additional functional validation are required to confirm this possibility. Likewise, a large comparison between clonal evolution patterns in *TP53* wild-type, null and mutated cases may help in the understanding of the different clinical outcome of these patients.

With respect to **endometrial carcinomas (ECs)**, little is known about ITH. The majority of the previous large-scale sequencing analyses in ECs have been performed on primary tumor samples and were focused on the molecular characterization of this type of tumors[188, 190-196]. In these studies, the classical dualistic model of EC dividing in endometrioid (EEC) and non-endometrioid (mostly represented by serous histology, SEC) subtypes[122], was complemented by the new molecular classification established by the TCGA study[196]. This classification includes POLE/ultramutated, MSI/hypermutated, copy-number low/endometrioid and copy-number high/serous-like subgroups. Most importantly, this categorization implies different clinical outcomes[122].

Regarding **ITH in tumor progression and EC evolution**, only one study has been published to date, in which atypical hyperplasia as a premalignant lesions, and paired samples of primary tumor and metastasis were analyzed[50]. In this study, multiregion analysis showed that metastatic samples were very homogeneous among them and usually arose from a common ancestral subclone that was not detected in the primary tumor biopsy, but it is important to mention that this work only included one sample from the primary tumor, which could be insufficient for understanding the phylogenetic evolution of EC.

Therefore, to have a wider and more precise picture of EC evolution, in this thesis we performed WES analysis including multiple regions not only from the metastatic disease but also from the primary tumor. Accordingly, we analyzed a total of 10 metastatic EC

patients that were histologically classified as EEC (n=7) and SEC (n=3), and molecularly subdivided regarding TCGA criteria[196] as copy-number low/endometrioid (EEC1-4), MSI/hypermutated (EEC5-7) and copy-number high/serous-like (SEC1-3). At least two regions from the primary tumor and one from the metastatic lesion were included in the WES, although additional tumor samples were further analyzed by targeted sequencing.

Our WES analysis revealed that less than half of mutations (44%) were shared between multiple samples of each EC patient, a result that is similar to the ITH observed in other tumor types like renal cancers[25, 27], but clearly different to the higher homogeneity found in pancreatic or lung carcinomas (where around 70% of the variants were ubiquitously detected)[24, 28, 29]. This suggests that single-region sequencing might not be sufficient in EC, because the majority of mutations would not be represented. Remarkably, different degrees of ITH between the multiple molecular subgroups were found, being these percentages significantly lower in the MSI subgroup (mean = 12%). Although this low frequency of ubiquitous mutations might be related to the high random mutational rate characteristic of MSI tumors, no evidence supporting this idea have been previously described. In fact, tumors with similar somatic mutation prevalence as small cell and non-small cell lung carcinomas showed high differences in the percentage of shared mutations (95% versus 50%, respectively)[16, 30]. Furthermore, the majority of detected variants were found in a clonal status in each sample, as was inferred by the calculation of the Cancer Cell Fraction (CCF)[43], suggesting that when a single sample was considered most of the cancer cells carried the majority of variants detected in this specific region. This could be misunderstood as an absence of heterogeneity. However, clonal mutations in one region of the tumor could be completely absent in another tumor region, as previously reported in renal[25], lung[28] or ovarian[113] cancers, among others.

Moreover, our study also revealed that the whole common SCNAs found among samples from the same tumor was very low (6%), although the percentage of shared genes affected by SCNAs was higher (21%) (**Figure 22**). These differences could be attributed to the fact that the genomic coordinates determined using the exome data are only an inference and might not be totally accurate[214]. Thus, these could show variations regarding specific SCNAs but including the same affected genes. Comparing our results to the Gibson *et al*. series[50], similar observations were found in terms of the percentage of shared mutations (44% versus 48%) but not in the common SCNAs (21% versus 56%). Nonetheless, it is important to remind that while our series included at least 3 samples per patient, Gibson *et al*.[50] used paired samples of a unique primary tumor region and a metastatic sample.

We next analyzed in detail the ITH at the SCNA level in the serous cancers, which are a molecular subgroup with extensive genomic alterations[196]. The three cases included in our

series showed more ITH at the SCNAs than at the genetic level (**Figure 23**). Interestingly, the metastatic samples showed an important fraction of the non-shared SCNAs, suggesting that big genomic changes may have an important role in the metastatic disease. These findings coincided with the previous observed in metastatic pancreatic tumors, where genomic instability increased during metastatic progression[26].

On the other hand, no differences were found in the number of genetic alterations when between primary tumor and metastatic lesions, and no clear evidences of metastasis-specific exomic mutations were either, as previously reported in endometrial and other cancer types[24, 50]. These results might indicate that genetic mutations do not have an essential role in metastasis development, or at least that there is not a common genetic mechanism to generate them in ECs. It is also important to note that other genomic alterations, as chromotripsis or big genomic changes, could be participating in the metastasis development as recently observed in pancreatic cancers[257], and should be further analyzed in the EC clinical context.

Concerning tumor progression, Gibson *et al*.[50] found a high proportion of monophyletic evolution in EC. In fact, in 6 of their 7 (86%) cases the metastatic lesions were more closely related to each other than to the primary tumor, which suggest a common evolution of the metastatic regions from an ancestral clone. By contrast, only one case showed a metastatic region closer at the mutational level to the primary tumor than to the other metastasic regions, as occur in polyphyletic tumor progression. However, in our series we observed monophyly in 7 of the 10 patients (70%) and polyphyletic evolution in the remaining (30%) analyzed cases (**Figure 24**). These discrepancies found in the phylogenetic patterns between both studies could be due to the absence of multiple regions from the primary tumor in the Gibson *et al*. study[50]. In fact, when additional tumoral samples were analyzed by targeted sequencing, the phylogenetic patterns were better clarified (**Figure 26-32**)**.** For example, the phylogenetic tree obtained for the EEC-1 patient (**Figure 26**) showed a clear polyphyletic evolution, being one metastatic lesion (M) genetically related to specific primary tumors regions (T4 and T2), and the other metastatic area (M2) close to T1 region. This seems to be a case of tumor with a complex metastatic seeding, known as oligoclonal evolution (**Figure 40a**), which had not been previously described in EC but is frequent in other tumor types as pancreatic[24] or prostate[40] tumors.

Nevertheless, the most common phylogenetic pattern found in our study was that in which metastatic lesions were closely related and arise from one branched subclone of the tumor (EEC-2, -3, -4, -6 and -7, and SEC-1-3). This homogeneity between metastatic samples was also found in the Gibson et al. series[50] as well as in pancreatic[26] and prostate[48] metastatic

tumors, and brain metastasis from lung, breast or renal carcinomas[49]. In our study, the subclone that leads to the metastasis could arise directly from a primary tumor region, as seems to occur in EEC-2-4 cases (**Figure 27-28, Figure 40b**), or from a common ancestor as in EEC-6-7 and SEC-1-3 patients (**Figure 30-32, Figure 40c**).

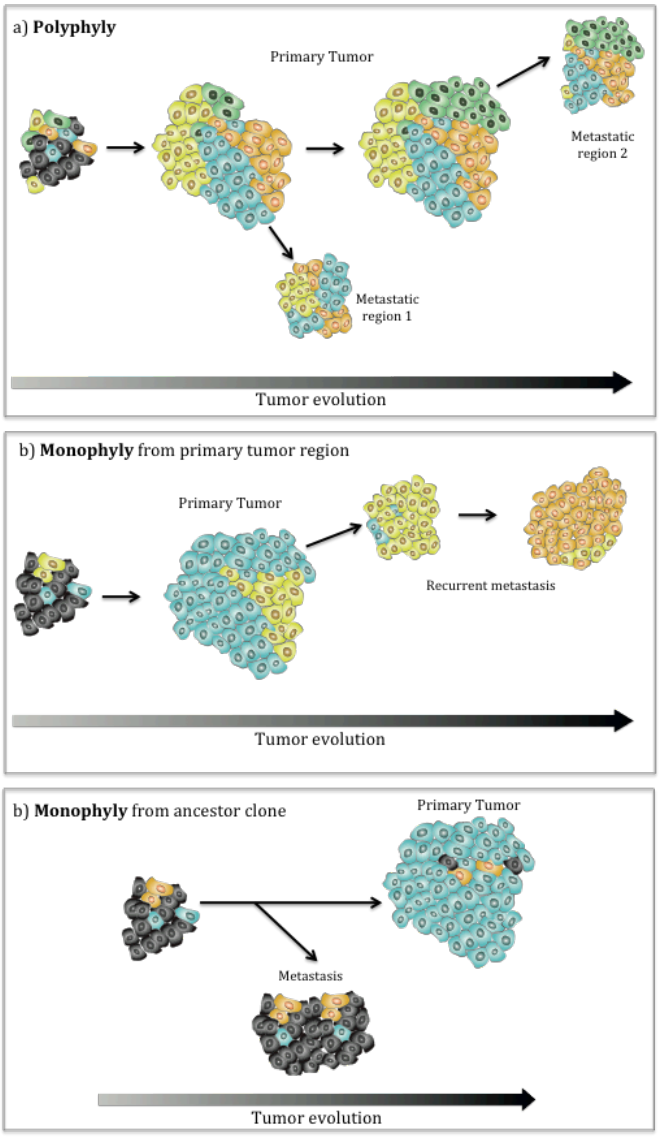


Figure 40 Next generation sequencing analyses revealed different phylogenetic evolution patterns in metastatic EC. WES followed by targeted massive parallel sequencing revealed different clonal evolution patterns in the metastatic EC analyzed. The most heterogeneous cases (EEC-1 and EEC-5) showed a polyphyletic evolution (a), while in those cases with less ITH monophyly was found. Two subtypes of monophyletic evolution patterns were identified, one of them with metastatic regions arising from primary tumor regions and mostly acquiring additional mutations in the metastasis (EEC-2-4) (b), and the other one with metastatic regions initially evolved from an ancestor clone and subsequent mutations acquired in the primary tumor (EEC-6-7 and SEC-1-3) (c).

Interestingly, in addition to the common mutations found in each tumor, the majority of the cases showed more specific variants in the primary tumor regions than in the metastatic lesions (EEC-2, -6, and -7 and SEC-1-3). In all these cases, except EEC-7 patient, the metastatic regions were detected at the moment of diagnosis and were located at the ovary. These results could be pointing out that the metastatic disease appear from an initial clone of the tumor with the specific molecular features to spread to the ovary, while the same tumor clone in the uterus would continue acquiring mutations across the time, modifying the initial tumor profile (**Figure 40c**). Remarkably, all the serous cases shared

this specific pattern and similar phylogenetic threes (**Figure 32**), though additional tumors should be analyzed in order to determine if this pattern could be related to tumor histology or to metastasis tropism. Similar results have been previously observed in synchronous endometrioid endometrial and ovarian carcinomas[209, 258]. In these studies, evidences of clonality were found between both tumors, indicating dissemination from one to other site. In those cases analyzed by WES, endometrial tumors showed more mutations than ovarian carcinomas in the majority of patients[209]. By contrast, recurrence regions were analyzed in the EEC-7 patient (**Figure 31**). In this case, although primary tumor and metastatic regions shared a low percentage of the variants, both developed exclusive changes in each group of samples. Interestingly, the number of specific mutations was higher in the primary tumor group of samples (T1-T3) than in the recurrences (M-M5). These results could suggest that independently of the particular mutational pattern found in the primary tumor and the relapse samples, the early progression of a clone would have given rise to the recurrent disease.

On the other hand, EEC-3 and EEC-4 patients showed more variants in their metastatic lesions, lymph nodes and epiplon respectively, than in the primary tumor regions (**Figure 28**). In both cases, metastases were detected as a recurrent disease after 3 years of the primary tumor treatment. These results may indicate that the recurrent metastatic samples acquire additional mutations during or after tumor treatment, which enabled their growth even after years of disease remission (**Figure 40b**). In these cases, metastasis would be slowly acquiring a more complex genetic phenotype that could led to an aggressive behavior during their development. These findings have been previously observed in ovarian[115, 117], pancreatic[24] and renal[25] carcinomas as well as in secondary acute myeloid leukemia[42].

As a whole, our study emphasizes the importance of analyzing the maximum number of samples to better understand tumor evolution, especially when the genetic heterogeneity found between the multiple samples of the primary tumor is considered. Nevertheless, it is fairly difficult to calculate the number of tumor regions needed to ensure proper assessment of ITH, since it is not possible to anticipate the ITH degree that each tumor will have. Another important point on this subject is the sequencing methodology used in the ITH study. In our project, part of the WES variants were re-analyzed by targeted sequencing. This technique is restricted to specific regions, which may limit overall tumor characterization, but allowed the study of additional FFPE tumor and metastatic samples. Moreover, this sequencing approach reaches coverage values more than 20-fold higher than WES, which was important in the sequencing of cases with unstable DNAs as those of MSI tumors. In these MSI tumor cases (EEC5-7), the percentage of shared mutations

significantly increases when compared WES and targeted sequencing (**Figure 25**), being more variants detected during the validation re-analysis. Similar results were found in the study of localized lung cancers, also characterized by a high somatic mutation prevalence[16, 29, 30]. Thus, targeted sequencing could generate more reliable data but limited to the sequenced regions.

In conclusion, different degrees of ITH as well as distinct patterns of tumor evolution were found in each EC patient, apparently independently of their histological or molecular classification. However, the relationship between ITH and prognosis or treatment response is yet to be determined. Although recent studies have suggested that ITH could be an independent prognostic factor of disease progression and survival in other types of tumors[43, 52], the mechanisms involved in this process remain still unknown. During the last years, the fact that some ECs remain resistant to treatment was hypothetically related with the presence of ITH in the primary tumor[156]. However, our data did not suggest any correlation between the heterogeneity found in the primary tumor and the presence of recurrent disease.

As a third aim, we evaluated the **potential use of WES in the diagnosis and genetic characterization of ambiguous pathological entities**. Some ECs, recently considered as morphological ambiguous[128, 129], remain controversial in their histological diagnosis, and in consequence in their clinical treatment and outcome. As this type of carcinomas is not considered a clinical entity nowadays, advancing in its characterization supposes a goal not only from a molecular point of view but also for its clinical management. Thus, until biologically and clinically validated diagnostic criteria are developed for these cases, the genomic characterization of these tumors could offer additional knowledge about its nature.

To this end, we performed WES in multiple samples from an ambiguous metastatic endometrial carcinoma (AEC) and its patient derived xenografts (PDXs). The histological classification of this case, based on the pathologist observations and the immunohistochemical profile of the tumor, showed a diagnostically challenging case that was finally classified as an ambiguous tumor (**Figure 33**). Moreover, after the bioinformatics analysis of WES results, the endometrium was confirmed as the tissue origin of this tumor (**Figure 34**). The WES analysis performed in two primary tumor regions and a metastatic lesion revealed the presence of ITH, being considerably higher at the SCNAs than at the genetic variants level (10% of shared genes affected by SCNAs versus 80% of shared genes with genetic mutations). Additionally, the calculation of the

cancer cell fraction (CCF) for each variant of the three samples revealed 80% of clonal mutations, indicating that the majority of cells of each sample carried most of the variants identified in this region (**Figure 35**). The similarity between the two primary tumor regions (T1 and T2) was higher than between them and metastatic lesion (M), supporting its potential monophyletic evolution. However, targeted sequencing of three additional tumor samples revealed a higher degree of ITH than that observed by WES, with only 17% of the analyzed variants common among these samples. Therefore, the phylogenetic tree showed a polyphyletic branched pattern of evolution, being the T4-T5-T3 tumor regions initial clones in the tumor progression, while the metastatic lesion was closely related to T1 and T2 regions (**Figure 36**). These controversial findings are again evidencing the necessity of including the maximum number of regions as well as using of high coverages during the sequencing study, in order to reliably understand the ITH of a tumor.

Surprisingly, none of the most frequently mutated genes in EC[50, 196] were detected in this tumor. In fact, only one of the identified variants was considered to have a real pathogenic impact in its corresponding protein. This variant, detected in all the analyzed samples, affected one splice acceptor site of the *Ataxia Telangiectasia Mutated* (*ATM*) gene, although its impact on protein function was not tested. On the other hand, the SCNAs study revealed a gain of material affecting the *CCND1* gene, which codifies for the cyclin D1 protein, thus suggesting a potential cell cycle malfunction in this tumor.

To obtain a broader view of the molecular pathways altered in this AEC, the mutational signature of all the samples was analyzed. It is known that diverse mutational processes produce different combinations of mutation types, called 'signatures' and numbered from 1 to 21 depending on their variation patterns[16]. In the three samples analyzed by WES in this ambiguous carcinoma, enrichment in the mutational signature 13 was found. This signature is produced by mutations generated by the APOBEC family of cytidine deaminases, which convert cytidine to uracil[259]. It is worth noting that this signature has not been found in ECs until now, but it has been observed in bladder and breast cancers[16]. Interestingly, it has been recently described the implication of DNA replication stress in the mutagenesis process mediated by the APOBEC3 enzymes, with the involvement of *ATM* and *CCND1* genes among others[260]. Altogether, these alterations could be implicated in the initial events of this tumor, maybe driving the mutational process and leading to this specific mutational signature. The results obtained for this tumor contrast with those obtained in the rest of metastatic ECs analyzed by WES, showing all of them mutational patterns typical of EC[16]. While the majority of CN-low tumors (EEC1-4) showed the mutational signature associated to the age of the patients (signature 1); as expected MSI cases (EEC5-7) were characterized by the mutational pattern related to DNA MMR

deficiency (signature 6). These results could be revealing that different mechanism underlying genetic instability give rise to several mutational signatures, being the AEC outside the general mechanism involved in ECs.

To investigate further about the biology of this ambiguous tumor, patient derived xenograft (PDX) models were generated from different regions of the primary tumor and metastatic locations. Nowadays, PDX models are popular preclinical models in translational oncology research due to their histological and genetic similarity to the original tumors, as has been reported for EC[261], and their potential use to predict therapeutic response to treatments[262].

In our study, a total of five PDX mice were obtained, representing regions from the superficial and deep primary tumor as well as areas from metastatic disease (right and left lymphatic nodes and cervix implants). Different genetic characterization approaches were performed in the multiple PDX samples with the aim of investigating its reliability representing the genomic landscape of this tumor patient. Firstly, a targeted massive parallel sequencing study revealed that the majority of mutations previously detected in the patient (79/86, 81%) were also found in all the analyzed PDXs, including that found in the *ATM* gene. However, most of the variants that were heterogeneously identified in the patient did not appear in the PDX samples. In this sense, our results could be suggesting that these PDX models did not completely represent the ITH found in the multiple tumor regions of the patient. To validate this hypothesis WES analysis was also performed in these samples, obtaining similar results than those observed in the targeted sequencing analysis. However, WES revealed that PDX tumors showed a high number of additional mutations not previously identified in the tumor patient. From the total of the PDX variants, around a half were those previously observed in the patient while the rest of them were new in the PDX samples (**Figure 37**). Altogether, these results suggested that the genetic profile identified in our PDX models significantly differs from the original found in this patient, and this is in striking contrast to similar studies with other tumor types, where PDX models mostly retained the molecular heterogeneity of their originating samples[263, 264]. This could indicate that, due to the original tumor of this patient did not show a high degree of heterogeneity, the minor subclones not represented in the PDX samples could not have the capability of growth in the PDX models. Besides, the additional mutations found in these PDX models may reflect a quick evolution of the tumor in the host, and could be related to selection pressure that occur during engraftment into diverse species (from human to mouse)[265, 266].

Despite the ITH differences between PDX and the patient tumor, we next tested the potential utility of these preclinical models for therapy selection, as has been previously

described in other tumor types[262]. We performed an *in silico* study based on the comparison of the mutational profile found in the AEC and the available information about the drugs related to this genetic status. Although this is a preliminary study, bortezomib and paclitaxel were identified as potentially effective treatment in this patient (**Figure 38a**). In this regard, the pathogenic mutation found in the *ATM* gene seems to be indicating a malfunction of the cell cycle and the DNA repair machinery that would favor the response of this tumor to these selected therapies[227, 267, 268]. It is important to notice that, although previous evidences related these genes with the correspondent treatment[227, 269, 270], functional analysis to verify the effect of each genetic variant would be needed. Regarding the selected treatments, bortezomib was the first proteasome-inhibitor to be approved in the clinics, and while it is mainly used in the treatment of multiple myeloma and mantel cell lymphoma[271, 272], multiple evidences indicate its potential utility in EC[273-277]. Paclitaxel is a well-known taxane that acts as an antimicrotubule agent, widely used in the treatment of multiple type of solid tumors, including EC[278].
In order to test the results obtained in the *in silico* study, PDX models generated from the AEC tumor were treated with the proposed therapies. Our preliminary data pointed out that the combination of paclitaxel with bortezomib showed a benefit in terms of tumor reduction compared to standard treatment protocol, although the differences were not statistically significant due to the low number of mice analyzed (**Figure 38b**). These results suggest the potential applicability of the predicted treatment in this patient and the utility of this *in silico* study, though it is worth mentioning that they have not been directly validated in the patient.

In summary, the combination of genomic profiling in tumors and PDX preclinical models would be potentially useful in the clinical practice to identify the most beneficial treatments for the patient, although they require testing large number of animals and treatment options, as well as a higher number of EC patients of diverse histological and molecular subtypes.

In addition to the clinical outcome and treatment response, another important problem derived from the ITH is the use of appropriate biomarkers for patient stratification[51]. During the last years, several clinical trials have moved from histological to molecular classification based on the somatic alterations of the tumors[279-281]. In this sense, in numerous clinical trials the patients with EC have been stratified according to their genetic tumor features, such as *PTEN*, *PIK3CA* or *FGFR3* mutational status, among others (https://clinicaltrials.gov)[156]. Moreover, a next generation of clinical trials (i.e., NCI-

MATCH, I-SPY[282], FOCUS or MATRIX) include multiple subgroups of tumor types treated with a specific targeted therapy. The use of proper samples to stratify the patients enrolled in these trials is essential and a current challenge in the molecular oncology. For example, archival material should not be used when recurrent disease is studied, underlining the need to re-biopse these cases. To this end, the Tracking Non-small Cell Lung Cancer Evolution Through Therapy (TRACERx) study has been designed with this aim, and includes multiple region samples as well as circulating DNA (ctDNA) and circulating tumor cells (CTCs)[283]. However, these kinds of approaches are nowadays unaffordable for the clinical application, and therefore, two different solutions have been proposed: a) combining DNA from different tumor samples before performing genetic analysis, but this would lead to a decrease in the frequency of subclonal mutations; b) the use of liquid biopsies, which could be the best option to capture ITH nowadays[53-55].

In this regard, another important objective of this thesis was focused on the **implementation of alternative biopsies in genetic characterization of EC to overcome ITH.**

Uterine aspirates (UAs) are low invasive and highly sensitive pre-operative biopsies[149], previously proposed as a molecular diagnostic tool on the basis of differential gene expression profiles[153, 154]. Thus, in the current study we evaluated the potential utility of UAs to uncover the overall tumor mutational profile by targeted massive parallel sequencing using a predesigned panel that includes the most frequently altered genes in EC (*PTEN*, *KRAS*, *FGFR2*, *CTNNB1*, *PIK3CA*, *FBXW7* and *TP53*). Consequently, it is a useful tool to characterize the mutational profile of endometrial tumors into the clinical practice since it considerably reduces the time of sample processing, data analysis as well as the cost of the procedure, compared to WES. Nevertheless, it is important to note that the panel used for this analysis is focused on point genetic mutations, less commonly found in SEC and CS, which are better characterized by big genomic changes[124].

Our study revealed that this approach allowed the identification of genetic variants in UAs not only in samples derived from tumor cases but also in those from atypical hyperplasia, while only 4% of the control patients without a malignant disease at the moment of the study showed mutations (**Figure 15**). Importantly, variants detected in the UAs accurately represent the molecular profile of their corresponding tumors (**Figure 16**), even in those UAs considered non-diagnosable by the pathologist. It has been described that around 13% of the UAs are not valid to distinguish the presence or absence of disease (non-diagnosable), either due to a small proportion of representative tumor cells or to the poor quality of the sample[146-148]. This percentage of cases is concordant with the fail rate found

in our series (8/62, 13%). Therefore, the fact that genetic sequencing of UAs produced valuable information even in non-diagnosable cases, confirms its potential utility in the pathological diagnosis in order to reduce its fail rate.

Moreover, the usefulness of UAs to recapitulate ITH in EC was also evaluated. Initially, a multiregion analysis was performed in EEC, SEC and CS samples, revealing that ITH was broadly present in EEC (71%) even when targeted sequencing is performed, but not so common in SEC or CS (20%) (**Figure 18** and **19**). These results showed that UAs detect a more representative mutational landscape than a unique surgical tumor specimen, reproducing in a single sample the ITH found in the multiple tumor regions. To better characterize this observation, a quality score defined as mutation discovery rate (MDR) was calculated as the percentage of mutations detected in each individual sample with respect to all mutations found in a given patient. Interestingly, MDR value was significantly higher in UAs than in the tumor regions from EECs, indicating that they are better biopsies to characterize the genetic profile of the tumor.

Regarding these important findings, we also analyzed the UAs from two of the metastatic EC included in the WES study (EEC2- and -3). In this case, genetic variants identified by WES were reanalyzed by targeted sequencing, being 72% (EEC-2) and 87% (EEC-3) of the analyzed variants also detected in the UA. Remarkably, in the UA study from EEC-3 case we observed mutations not identified in the primary tumor regions but mainly detected in the metastatic samples, which arised from recurrent disease diagnosed three years after primary tumor diagnosis. In this regard, the study of the aspirate could be representing even those mutations present in tumor subclones that would progress and evolve years later. As this biopsy was obtained during the diagnosis of the primary tumor, these results would be indicating the presence of this mutational pattern before treatment, as well as the capability of UAs to represent the majority of the subclones that make up the tumor.

These results highlight the powerful utility of this type of alternative biopsy in combination with its molecular characterization and confirmed that the use of a unique tumor sample in diagnosis could underestimate the mutational landscape in heterogeneous tumors. The fact that ITH is well represented in aspirates is probably related to the nature of such samples, consisting of cells from many different parts of the uterine cavity, thus providing a more representative picture of the entire tumor specimen than a single sample from a specific tumor region (**Figure 41**). Similar results were observed in ovarian carcinomas, where ITH was represented in the liquid ascites[114, 284]. In this case, ascites could be representing the entire cavity in a similar way that UAs do in uterine cancer, capturing all the genetic alterations and representing the ITH found in the

different tissue tumor samples. This underlines the importance of further investigating the application of other organic fluids and its diagnostic implementation, for example, peritoneal lavages in gastrointestinal tumors or fine needle aspiration cytology for pulmonary malignancies. However, the use of non-invasive biopsies[53-55] easily obtained from a routine blood tests, as ctDNA or CTCs are recommended not only to unravel ITH but also to facilitate patient follow up in the near future[53, 55, 285]. In this regard, the use of CTCs as a prognosis marker in EC has been previously described[286], and its genetic characterization could be helpful during tumor development.

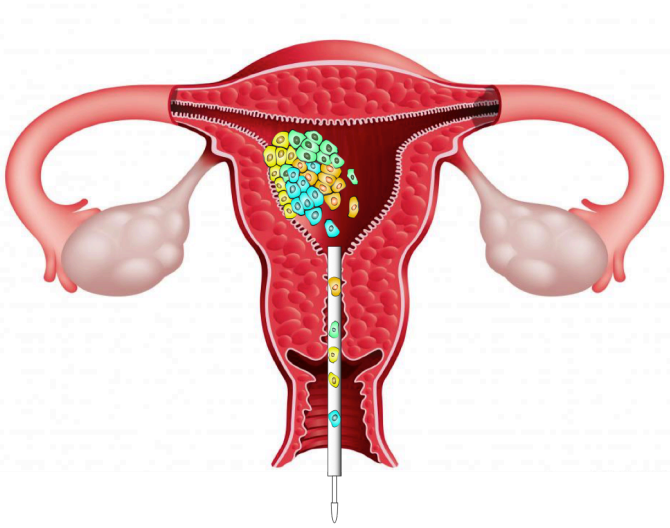
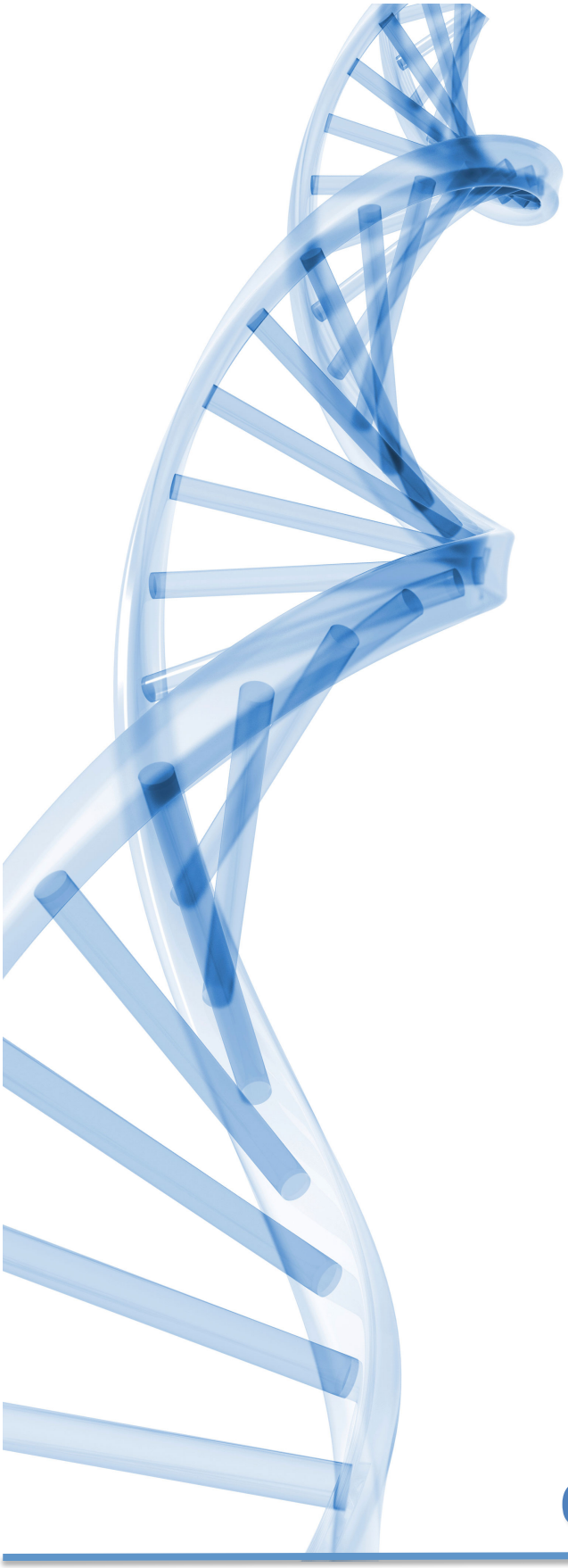


Figure 41 Uterine aspirates represent ITH found in the tumor multiregion analysis. Genetic analysis revealed that uterine aspirates capture intratumor heterogeneity found in the different regions of the patient tumor (represented with different colored cells), probably due to the fact that uterine aspirates are samples that contain cell from all different parts of the uterine cavity.

Finally and to summarize, throughout this thesis multiple approaches as well as tumor types were analyzed in order to decipher the implication of ITH in gynecological cancers. The results showed in this work have shed light on to the clonal evolution and the role of ITH in gynecological cancers, as well as the potential use of pre-operative diagnostic biopsies in this specific clinical context. Nevertheless, an in-depth study considering a larger number of EC patients should be undertaken, thus, there is a long way to go in the clinical application of these findings.



CONCLUSIONS

6 CONCLUSIONS

1. Massive (whole-exome and targeted) sequencing analyses allow the characterization of different spatial and/or temporal intratumor heterogeneity degrees in multiple subtypes of gynecological cancer, including a *TP53*-null high grade serous ovarian carcinoma and several metastatic and non-metastatic endometrial carcinomas.

2. Functional analysis based on the mutational profile of a patient may reveal signaling pathways involved in tumorigenesis and might be use as a complementary tool to reach a global tumor understanding.

3. The analysis of the maximum number of regions per tumor as well as high sequencing coverage values achieve a better intratumor heterogeneity and clonal evolution characterization.

4. Different phylogenetic evolution and clonality patterns were found in metastatic endometrial carcinomas, apparently not related to the classical histological or molecular classification.

5. The absence of additional mutations in metastatic regions compared to the primary tumor could be suggesting their irrelevant role during endometrial tumor progression.

6. Genomic and genetic characterization of ambiguous endometrial carcinomas could unravel their molecular features and probably help in their clinical management.

7. Patient derived xenograft (PDX) genetic profile retains the majority of common mutations present in the whole human tumor. Nevertheless, additional variants found in the mice could be indicating a new tumor evolution within the host.

8. Genetic analysis in uterine aspirates recapitulates the mutational profile of endometrial carcinomas, even in non-diagnosable cases, and captures intratumor heterogeneity detected in the multiple endometrial tumor regions.

CONCLUSIONES

1. Los análisis de secuenciación masiva (tanto exomas como dirigida) permiten la caracterización de la heterogeneidad intratumoral espacial y temporal en cáncer ginecológico, incluyendo en un tumor seroso de ovario de alto grado con mutación nula en TP53 así como en varios tipos de carcinomas de endometrio.

2. Los análisis funcionales basados en el perfil mutacional de un paciente pueden generar información sobre de las vías de señalización implicadas en su tumorogénesis, y tienen utilidad como una herramienta complementaria para entender la biología del tumor.

3. El estudio mutacional de un elevado número de regiones del tumor así como el uso de altos valores de cobertura en la secuenciación masiva permiten caracterizar de una forma fiable tanto la evolución clonal como la heterogeneidad intratumoral presente.

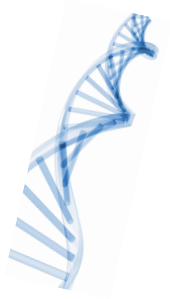
4. De acuerdo al análisis de secuenciación masiva se han podido identificar diferentes patrones de evolución filogenética y composición clonal en los carcinomas de endometrio metastásicos, sin encontrar una relación aparente con la clasificación histológica y molecular clásica en estos tumores.

5. La ausencia de mutaciones adicionales en las regiones metastásicas de carcinomas de endometrio en comparación con el tumor primario sugiere que dichas alteraciones no están implicadas en la progresión tumoral.

6. El uso de aproximaciones basadas en el análisis genómico de carcinomas ambiguos de endometrio puede ayudar a mejorar su manejo clínico a través de su caracterización molecular.

7. El perfil genético de los modelos animales PDX (patient derived xenograft) refleja la mayoría de las mutaciones presentes en el tumor de origen. Sin embargo, la presencia de variantes genéticas adicionales en los PDX podrían sugerir una nueva evolución tumoral en el huésped.

8. El análisis genético de los aspirados uterinos representa el perfil mutacional de los carcinomas de endometrio de los que proceden, incluso en casos de difícil diagnóstico, y captura la heterogeneidad intratumoral detectada en las múltiples muestras del tumor.



BIBLIOGRAPHY

7 BIBLIOGRAPHY

1. Dulbecco, R., *A turning point in cancer research: sequencing the human genome.* Science, 1986. **231**(4742): p. 1055-6.
2. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
3. Venter, J.C., et al., *The sequence of the human genome.* Science, 2001. **291**(5507): p. 1304-51.
4. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
5. Futreal, P.A., et al., *A census of human cancer genes.* Nat Rev Cancer, 2004. **4**(3): p. 177-83.
6. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**(7057): p. 376-80.
7. Shendure, J., et al., *Accurate multiplex polony sequencing of an evolved bacterial genome.* Science, 2005. **309**(5741): p. 1728-32.
8. Ley, T.J., et al., *DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome.* Nature, 2008. **456**(7218): p. 66-72.
9. Wheeler, D.A., et al., *The complete genome of an individual by massively parallel DNA sequencing.* Nature, 2008. **452**(7189): p. 872-6.
10. Wood, L.D., et al., *The genomic landscapes of human breast and colorectal cancers.* Science, 2007. **318**(5853): p. 1108-13.
11. Cancer Genome Atlas Research, N., et al., *The Cancer Genome Atlas Pan-Cancer analysis project.* Nat Genet, 2013. **45**(10): p. 1113-20.
12. Wheeler, D.A. and L. Wang, *From human genome to cancer genome: the first decade.* Genome Res, 2013. **23**(7): p. 1054-62.
13. Lawrence, M.S., et al., *Discovery and saturation analysis of cancer genes across 21 tumour types.* Nature, 2014. **505**(7484): p. 495-501.
14. Chmielecki, J. and M. Meyerson, *DNA sequencing of cancer: what have we learned?* Annu Rev Med, 2014. **65**: p. 63-79.
15. Huang, F.W., et al., *Highly recurrent TERT promoter mutations in human melanoma.* Science, 2013. **339**(6122): p. 957-9.
16. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer.* Nature, 2013. **500**(7463): p. 415-21.
17. Shen, T., et al., *Clinical applications of next generation sequencing in cancer: from panels, to exomes, to genomes.* Front Genet, 2015. **6**: p. 215.
18. Park, J.Y., et al., *Clinical genomics: when whole genome sequencing is like a whole-body CT scan.* Clin Chem, 2014. **60**(11): p. 1390-2.
19. Luthra, R., et al., *Next-Generation Sequencing in Clinical Molecular Diagnostics of Cancer: Advantages and Challenges.* Cancers (Basel), 2015. **7**(4): p. 2023-36.
20. Nowell, P.C., *The clonal evolution of tumor cell populations.* Science, 1976. **194**(4260): p. 23-8.
21. Spremulli, E.N. and D.L. Dexter, *Human tumor cell heterogeneity and metastasis.* J Clin Oncol, 1983. **1**(8): p. 496-509.
22. Shackney, S.E. and T.V. Shankey, *Genetic and phenotypic heterogeneity of human malignancies: finding order in chaos.* Cytometry, 1995. **21**(1): p. 2-5.
23. Navin, N., et al., *Inferring tumor progression from genomic heterogeneity.* Genome Res, 2010. **20**(1): p. 68-80.
24. Yachida, S., et al., *Distant metastasis occurs late during the genetic evolution of pancreatic cancer.* Nature, 2010. **467**(7319): p. 1114-7.

25. Gerlinger, M., et al., *Intratumor heterogeneity and branched evolution revealed by multiregion sequencing.* N Engl J Med, 2012. **366**(10): p. 883-92.
26. Campbell, P.J., et al., *The patterns and dynamics of genomic instability in metastatic pancreatic cancer.* Nature, 2010. **467**(7319): p. 1109-13.
27. Gerlinger, M., et al., *Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing.* Nat Genet, 2014. **46**(3): p. 225-33.
28. de Bruin, E.C., et al., *Spatial and temporal diversity in genomic instability processes defines lung cancer evolution.* Science, 2014. **346**(6206): p. 251-6.
29. Zhang, J., et al., *Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing.* Science, 2014. **346**(6206): p. 256-9.
30. Saber, A., et al., *Mutation patterns in small cell and non-small cell lung cancer patients suggest a different level of heterogeneity between primary and metastatic tumors.* Carcinogenesis, 2016.
31. Nik-Zainal, S., et al., *The life history of 21 breast cancers.* Cell, 2012. **149**(5): p. 994-1007.
32. Shah, S.P., et al., *The clonal and mutational evolution spectrum of primary triple-negative breast cancers.* Nature, 2012. **486**(7403): p. 395-9.
33. Demeulemeester, J., et al., *Tracing the origin of disseminated tumor cells in breast cancer using single-cell sequencing.* Genome Biol, 2016. **17**(1): p. 250.
34. Thirlwell, C., et al., *Clonality assessment and clonal ordering of individual neoplastic crypts shows polyclonality of colorectal adenomas.* Gastroenterology, 2010. **138**(4): p. 1441-54, 1454 e1-7.
35. Wu, H., et al., *Evolution and heterogeneity of non-hereditary colorectal cancer revealed by single-cell exome sequencing.* Oncogene, 2016.
36. Sottoriva, A., et al., *Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics.* Proc Natl Acad Sci U S A, 2013. **110**(10): p. 4009-14.
37. Kumar, A., et al., *Deep sequencing of multiple regions of glial tumors reveals spatial heterogeneity for mutations in clinically relevant genes.* Genome Biol, 2014. **15**(12): p. 530.
38. Johnson, B.E., et al., *Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma.* Science, 2014. **343**(6167): p. 189-93.
39. Haffner, M.C., et al., *Tracking the clonal origin of lethal prostate cancer.* J Clin Invest, 2013. **123**(11): p. 4918-22.
40. Gundem, G., et al., *The evolutionary history of lethal metastatic prostate cancer.* Nature, 2015. **520**(7547): p. 353-7.
41. Anderson, K., et al., *Genetic variegation of clonal architecture and propagating cells in leukaemia.* Nature, 2011. **469**(7330): p. 356-61.
42. Walter, M.J., et al., *Clonal architecture of secondary acute myeloid leukemia.* N Engl J Med, 2012. **366**(12): p. 1090-8.
43. Landau, D.A., et al., *Evolution and impact of subclonal mutations in chronic lymphocytic leukemia.* Cell, 2013. **152**(4): p. 714-26.
44. Bianchi, G. and I.M. Ghobrial, *Biological and Clinical Implications of Clonal Heterogeneity and Clonal Evolution in Multiple Myeloma.* Curr Cancer Ther Rev, 2014. **10**(2): p. 70-79.
45. Hughes, A.E., et al., *Clonal architecture of secondary acute myeloid leukemia defined by single-cell sequencing.* PLoS Genet, 2014. **10**(7): p. e1004462.
46. Bolli, N., et al., *Heterogeneity of genomic evolution and mutational profiles in multiple myeloma.* Nat Commun, 2014. **5**: p. 2997.
47. Evans, T. and U. Matulonis, *Next-Generation Sequencing: Role in Gynecologic Cancers.* J Natl Compr Canc Netw, 2016. **14**(9): p. 1165-73.
48. Liu, W., et al., *Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer.* Nat Med, 2009. **15**(5): p. 559-65.

49. Brastianos, P.K., et al., *Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic Targets.* Cancer Discov, 2015. **5**(11): p. 1164-77.
50. Gibson, W.J., et al., *The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis.* Nat Genet, 2016. **48**(8): p. 848-55.
51. Jamal-Hanjani, M., et al., *Translational implications of tumor heterogeneity.* Clin Cancer Res, 2015. **21**(6): p. 1258-66.
52. Morris, L.G., et al., *Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival.* Oncotarget, 2016. **7**(9): p. 10051-63.
53. De Mattos-Arruda, L., et al., *Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle.* Ann Oncol, 2014. **25**(9): p. 1729-35.
54. De Mattos-Arruda, L., et al., *Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma.* Nat Commun, 2015. **6**: p. 8839.
55. Jamal-Hanjani, M., et al., *Detection of ubiquitous and heterogeneous mutations in cell-free DNA from patients with early-stage non-small-cell lung cancer.* Ann Oncol, 2016. **27**(5): p. 862-7.
56. Latarjet M., R.L.A., Pró E., *Anatomía Humana*, ed. Panamericana. 2004, Barcelona.
57. WHO, *WHO Classification of Tumours of Female Reproductive Organs*. Vol. 6. 2014.
58. Ferlay J, S.I., Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11* Lyon, France: International Agency for Research on Cancer 2013.
59. AECC. *Website:* https://www.aecc.es/. Asociacion Española Contra el Cancer 2016.
60. Erickson, B.K., M.G. Conner, and C.N. Landen, Jr., *The role of the fallopian tube in the origin of ovarian cancer.* Am J Obstet Gynecol, 2013. **209**(5): p. 409-14.
61. Prat, J., *Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features.* Virchows Arch, 2012. **460**(3): p. 237-49.
62. Koonings, P.P., et al., *Relative frequency of primary ovarian neoplasms: a 10-year review.* Obstet Gynecol, 1989. **74**(6): p. 921-6.
63. Aune, D., et al., *Anthropometric factors and ovarian cancer risk: a systematic review and nonlinear dose-response meta-analysis of prospective studies.* Int J Cancer, 2015. **136**(8): p. 1888-98.
64. Beral, V., et al., *Ovarian cancer and hormone replacement therapy in the Million Women Study.* Lancet, 2007. **369**(9574): p. 1703-10.
65. Kim, H.S., et al., *Risk and prognosis of ovarian cancer in women with endometriosis: a meta-analysis.* Br J Cancer, 2014. **110**(7): p. 1878-90.
66. Lee, J.Y., et al., *Diabetes mellitus and ovarian cancer risk: a systematic review and meta-analysis of observational studies.* Int J Gynecol Cancer, 2013. **23**(3): p. 402-12.
67. Schildkraut, J.M., N. Risch, and W.D. Thompson, *Evaluating genetic association among ovarian, breast, and endometrial cancer: evidence for a breast/ovarian cancer relationship.* Am J Hum Genet, 1989. **45**(4): p. 521-9.
68. Collaborative Group on Epidemiological Studies of Ovarian, C., et al., *Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23,257 women with ovarian cancer and 87,303 controls.* Lancet, 2008. **371**(9609): p. 303-14.
69. Gong, T.T., et al., *Age at menarche and risk of ovarian cancer: a meta-analysis of epidemiological studies.* Int J Cancer, 2013. **132**(12): p. 2894-900.
70. Rice, M.S., M.A. Murphy, and S.S. Tworoger, *Tubal ligation, hysterectomy and ovarian cancer: A meta-analysis.* J Ovarian Res, 2012. **5**(1): p. 13.
71. Adami, H.O., et al., *Parity, age at first childbirth, and risk of ovarian cancer.* Lancet, 1994. **344**(8932): p. 1250-4.

72. Luan, N.N., et al., *Breastfeeding and ovarian cancer risk: a meta-analysis of epidemiologic studies.* Am J Clin Nutr, 2013. **98**(4): p. 1020-31.
73. Pal, T., et al., *BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases.* Cancer, 2005. **104**(12): p. 2807-16.
74. Risch, H.A., et al., *Population BRCA1 and BRCA2 mutation frequencies and cancer penetrances: a kin-cohort study in Ontario, Canada.* J Natl Cancer Inst, 2006. **98**(23): p. 1694-706.
75. Watson, P., et al., *The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome.* Int J Cancer, 2008. **123**(2): p. 444-9.
76. Bottoni, P. and R. Scatena, *The Role of CA 125 as Tumor Marker: Biochemical and Clinical Aspects.* Adv Exp Med Biol, 2015. **867**: p. 229-44.
77. Bast, R.C., Jr., et al., *CA 125: the past and the future.* Int J Biol Markers, 1998. **13**(4): p. 179-87.
78. Sarandakou, A., E. Protonotariou, and D. Rizos, *Tumor markers in biological fluids associated with pregnancy.* Crit Rev Clin Lab Sci, 2007. **44**(2): p. 151-78.
79. Hirsch, M., et al., *Diagnostic accuracy of cancer antigen 125 for endometriosis: a systematic review and meta-analysis.* BJOG, 2016. **123**(11): p. 1761-8.
80. Kipps, E., D.S. Tan, and S.B. Kaye, *Meeting the challenge of ascites in ovarian cancer: new avenues for therapy and research.* Nat Rev Cancer, 2013. **13**(4): p. 273-82.
81. Webber, K. and M. Friedlander, *Chemotherapy for epithelial ovarian, fallopian tube and primary peritoneal cancer.* Best Pract Res Clin Obstet Gynaecol, 2016.
82. Miller, D.S., et al., *Phase II evaluation of pemetrexed in the treatment of recurrent or persistent platinum-resistant ovarian or primary peritoneal carcinoma: a study of the Gynecologic Oncology Group.* J Clin Oncol, 2009. **27**(16): p. 2686-91.
83. Piccart, M.J., H. Lamb, and J.B. Vermorken, *Current and future potential roles of the platinum drugs in the treatment of ovarian cancer.* Ann Oncol, 2001. **12**(9): p. 1195-203.
84. Wiedemeyer, W.R., J.A. Beach, and B.Y. Karlan, *Reversing Platinum Resistance in High-Grade Serous Ovarian Carcinoma: Targeting BRCA and the Homologous Recombination System.* Front Oncol, 2014. **4**: p. 34.
85. Konecny, G.E. and R.S. Kristeleit, *PARP inhibitors for BRCA1/2-mutated and sporadic ovarian cancer: current practice and future directions.* Br J Cancer, 2016.
86. Nijman, S.M., *Synthetic lethality: general principles, utility and detection using genetic screens in human cells.* FEBS Lett, 2011. **585**(1): p. 1-6.
87. Ledermann, J.A., Y. Drew, and R.S. Kristeleit, *Homologous recombination deficiency and ovarian cancer.* Eur J Cancer, 2016. **60**: p. 49-58.
88. McClung, E.C. and R.M. Wenham, *Profile of bevacizumab in the treatment of platinum-resistant ovarian cancer: current perspectives.* Int J Womens Health, 2016. **8**: p. 59-75.
89. Ahmed, A.A., et al., *Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary.* J Pathol, 2010. **221**(1): p. 49-56.
90. Cancer Genome Atlas Research, N., *Integrated genomic analyses of ovarian carcinoma.* Nature, 2011. **474**(7353): p. 609-15.
91. Vogelstein, B., D. Lane, and A.J. Levine, *Surfing the p53 network.* Nature, 2000. **408**(6810): p. 307-10.
92. Bieging, K.T., S.S. Mello, and L.D. Attardi, *Unravelling mechanisms of p53-mediated tumour suppression.* Nat Rev Cancer, 2014. **14**(5): p. 359-70.
93. Bouaoun, L., et al., *TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data.* Hum Mutat, 2016. **37**(9): p. 865-76.
94. Kobel, M., et al., *The biological and clinical value of p53 expression in pelvic high-grade serous carcinomas.* J Pathol, 2010. **222**(2): p. 191-8.
95. Acs, G., T. Pasha, and P.J. Zhang, *WT1 is differentially expressed in serous, endometrioid, clear cell, and mucinous carcinomas of the peritoneum, fallopian tube, ovary, and endometrium.* Int J Gynecol Pathol, 2004. **23**(2): p. 110-8.

96. O'Neill, C.J. and W.G. McCluggage, *p16 expression in the female genital tract and its value in diagnosis.* Adv Anat Pathol, 2006. **13**(1): p. 8-15.
97. Esteller, M., et al., *Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors.* J Natl Cancer Inst, 2000. **92**(7): p. 564-9.
98. Walsh, T., et al., *Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing.* Proc Natl Acad Sci U S A, 2011. **108**(44): p. 18032-7.
99. Bowtell, D.D., *The genesis and evolution of high-grade serous ovarian cancer.* Nat Rev Cancer, 2010. **10**(11): p. 803-8.
100. Baker, V.V., et al., *c-myc amplification in ovarian cancer.* Gynecol Oncol, 1990. **38**(3): p. 340-2.
101. Slamon, D.J., et al., *Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer.* Science, 1989. **244**(4905): p. 707-12.
102. Muller, P.A. and K.H. Vousden, *p53 mutations in cancer.* Nat Cell Biol, 2013. **15**(1): p. 2-8.
103. Robles, A.I. and C.C. Harris, *Clinical outcomes and correlates of TP53 mutations and cancer.* Cold Spring Harb Perspect Biol, 2010. **2**(3): p. a001016.
104. Hall, J., J. Paul, and R. Brown, *Critical evaluation of p53 as a prognostic marker in ovarian cancer.* Expert Rev Mol Med, 2004. **6**(12): p. 1-20.
105. Shahin, M.S., et al., *The prognostic significance of p53 tumor suppressor gene alterations in ovarian carcinoma.* Cancer, 2000. **89**(9): p. 2006-17.
106. Wojnarowicz, P.M., et al., *The genomic landscape of TP53 and p53 annotated high grade ovarian serous carcinomas from a defined founder population associated with patient outcome.* PLoS One, 2012. **7**(9): p. e45484.
107. Petitjean, A., et al., *TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes.* Oncogene, 2007. **26**(15): p. 2157-65.
108. Wong, K.K., et al., *Poor survival with wild-type TP53 ovarian cancer?* Gynecol Oncol, 2013. **130**(3): p. 565-9.
109. Patch, A.M., et al., *Whole-genome characterization of chemoresistant ovarian cancer.* Nature, 2015. **521**(7553): p. 489-94.
110. Khalique, L., et al., *Genetic intra-tumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours.* J Pathol, 2007. **211**(3): p. 286-95.
111. Khalique, L., et al., *The clonal evolution of metastases from primary serous epithelial ovarian cancers.* Int J Cancer, 2009. **124**(7): p. 1579-86.
112. Cooke, S.L., et al., *Genomic analysis of genetic heterogeneity and evolution in high-grade serous ovarian carcinoma.* Oncogene, 2010. **29**(35): p. 4905-13.
113. Bashashati, A., et al., *Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling.* J Pathol, 2013. **231**(1): p. 21-34.
114. Castellarin, M., et al., *Clonal evolution of high-grade serous ovarian carcinoma from primary to recurrent disease.* J Pathol, 2013. **229**(4): p. 515-24.
115. Schwarz, R.F., et al., *Spatial and temporal heterogeneity in high-grade serous ovarian cancer: a phylogenetic analysis.* PLoS Med, 2015. **12**(2): p. e1001789.
116. Choi, Y.J., et al., *Intra-individual genomic heterogeneity of high-grade serous carcinoma of ovary and clinical utility of ascitic cancer cells for mutation profiling.* J Pathol, 2016.
117. Lambrechts, S., et al., *Genetic heterogeneity after first-line chemotherapy in high-grade serous ovarian cancer.* Eur J Cancer, 2016. **53**: p. 51-64.
118. Hoogstraat, M., et al., *Genomic and transcriptomic plasticity in treatment-naive ovarian cancer.* Genome Res, 2014. **24**(2): p. 200-11.
119. Huang, A., et al., *Detecting Circulating Tumor DNA in Hepatocellular Carcinoma Patients Using Droplet Digital PCR Is Feasible and Reflects Intratumoral Heterogeneity.* J Cancer, 2016. **7**(13): p. 1907-1914.

120. Sankaranarayanan, R. and J. Ferlay, *Worldwide burden of gynaecological cancer: the size of the problem.* Best Pract Res Clin Obstet Gynaecol, 2006. **20**(2): p. 207-25.
121. Bokhman, J.V., *Two pathogenetic types of endometrial carcinoma.* Gynecol Oncol, 1983. **15**(1): p. 10-7.
122. Lax, S.F. and R.J. Kurman, *A dualistic model for endometrial carcinogenesis based on immunohistochemical and molecular genetic analyses.* Verh Dtsch Ges Pathol, 1997. **81**: p. 228-32.
123. Murali, R., R.A. Soslow, and B. Weigelt, *Classification of endometrial carcinoma: more than two types.* Lancet Oncol, 2014. **15**(7): p. e268-78.
124. Yeramian, A., et al., *Endometrial carcinoma: molecular alterations involved in tumor development and progression.* Oncogene, 2013. **32**(4): p. 403-13.
125. Voss, M.A., et al., *Should grade 3 endometrioid endometrial carcinoma be considered a type 2 cancer-a clinical and pathological evaluation.* Gynecol Oncol, 2012. **124**(1): p. 15-20.
126. Cantrell, L.A., S.V. Blank, and L.R. Duska, *Uterine carcinosarcoma: A review of the literature.* Gynecol Oncol, 2015. **137**(3): p. 581-8.
127. Artioli, G., et al., *Rare uterine cancer: carcinosarcomas. Review from histology to treatment.* Crit Rev Oncol Hematol, 2015. **94**(1): p. 98-104.
128. Soslow, R.A., *Endometrial carcinomas with ambiguous features.* Semin Diagn Pathol, 2010. **27**(4): p. 261-73.
129. Espinosa, I., et al., *Mixed and Ambiguous Endometrial Carcinomas: A Heterogenous Group of Tumors With Different Clinicopathologic and Molecular Genetic Features.* Am J Surg Pathol, 2016. **40**(7): p. 972-81.
130. Zhang, Y., et al., *Overweight, obesity and endometrial cancer risk: results from a systematic review and meta-analysis.* Int J Biol Markers, 2014. **29**(1): p. e21-9.
131. Wise, M.R., et al., *Obesity and endometrial hyperplasia and cancer in premenopausal women: A systematic review.* Am J Obstet Gynecol, 2016. **214**(6): p. 689 e1-689 e17.
132. Allen, N.E., et al., *Endogenous sex hormones and endometrial cancer risk in women in the European Prospective Investigation into Cancer and Nutrition (EPIC).* Endocr Relat Cancer, 2008. **15**(2): p. 485-97.
133. Beral, V., et al., *Endometrial cancer and hormone-replacement therapy in the Million Women Study.* Lancet, 2005. **365**(9470): p. 1543-51.
134. Dossus, L., et al., *Reproductive risk factors and endometrial cancer: the European Prospective Investigation into Cancer and Nutrition.* Int J Cancer, 2010. **127**(2): p. 442-51.
135. Luo, J., et al., *Association between diabetes, diabetes treatment and risk of developing endometrial cancer.* Br J Cancer, 2014. **111**(7): p. 1432-9.
136. Shafiee, M.N., et al., *Reviewing the molecular mechanisms which increase endometrial cancer (EC) risk in women with polycystic ovarian syndrome (PCOS): time for paradigm shift?* Gynecol Oncol, 2013. **131**(2): p. 489-92.
137. Hemminki, K., L. Aaltonen, and X. Li, *Subsequent primary malignancies after endometrial carcinoma and ovarian carcinoma.* Cancer, 2003. **97**(10): p. 2432-9.
138. Druesne-Pecollo, N., et al., *Excess body weight and second primary cancer risk after breast cancer: a systematic review and meta-analysis of prospective studies.* Breast Cancer Res Treat, 2012. **135**(3): p. 647-54.
139. Jones, M.E., et al., *Endometrial cancer survival after breast cancer in relation to tamoxifen treatment: pooled results from three countries.* Breast Cancer Res, 2012. **14**(3): p. R91.
140. Mueck, A.O., H. Seeger, and T. Rabe, *Hormonal contraception and risk of endometrial cancer: a systematic review.* Endocr Relat Cancer, 2010. **17**(4): p. R263-71.
141. Schonfeld, S.J., et al., *An aggregated analysis of hormonal factors and endometrial cancer risk by parity.* Cancer, 2013. **119**(7): p. 1393-401.
142. Schmid, D., et al., *A systematic review and meta-analysis of physical activity and endometrial cancer risk.* Eur J Epidemiol, 2015. **30**(5): p. 397-412.

143. Mencaglia, L., et al., *Endometrial carcinoma and its precursors: early detection and treatment.* Int J Gynaecol Obstet, 1990. **31**(2): p. 107-16.
144. Bettocchi, S., et al., *Diagnostic inadequacy of dilatation and curettage.* Fertil Steril, 2001. **75**(4): p. 803-5.
145. Ceci, O., et al., *Comparison of hysteroscopic and hysterectomy findings for assessing the diagnostic accuracy of office hysteroscopy.* Fertil Steril, 2002. **78**(3): p. 628-31.
146. McCluggage, W.G., *My approach to the interpretation of endometrial biopsies and curettings.* J Clin Pathol, 2006. **59**(8): p. 801-12.
147. Phillips, V. and W.G. McCluggage, *Results of a questionnaire regarding criteria for adequacy of endometrial biopsies.* J Clin Pathol, 2005. **58**(4): p. 417-9.
148. Kandil, D., et al., *Clinical outcomes of patients with insufficient sample from endometrial biopsy or curettage.* Int J Gynecol Pathol, 2014. **33**(5): p. 500-6.
149. Clark, T.J., et al., *Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: a systematic quantitative review.* BJOG, 2002. **109**(3): p. 313-21.
150. Guido, R.S., et al., *Pipelle endometrial sampling. Sensitivity in the detection of endometrial cancer.* J Reprod Med, 1995. **40**(8): p. 553-5.
151. Bradley, L.D., *Complications in hysteroscopy: prevention, treatment and legal risk.* Curr Opin Obstet Gynecol, 2002. **14**(4): p. 409-15.
152. Chang, Y.N., et al., *Effect of hysteroscopy on the peritoneal dissemination of endometrial cancer cells: a meta-analysis.* Fertil Steril, 2011. **96**(4): p. 957-61.
153. Perez-Sanchez, C., et al., *Molecular diagnosis of endometrial cancer from uterine aspirates.* Int J Cancer, 2013. **133**(10): p. 2383-91.
154. Colas, E., et al., *Molecular markers of endometrial carcinoma detected in uterine aspirates.* Int J Cancer, 2011. **129**(10): p. 2435-44.
155. Dizon, D.S., *Treatment options for advanced endometrial carcinoma.* Gynecol Oncol, 2010. **117**(2): p. 373-81.
156. Eritja, N., et al., *Endometrial Carcinoma: Specific Targeted Pathways.* Adv Exp Med Biol, 2017. **943**: p. 149-207.
157. Gehrig, P.A. and V.L. Bae-Jump, *Promising novel therapies for the treatment of endometrial cancer.* Gynecol Oncol, 2010. **116**(2): p. 187-94.
158. Fleming, G.F., et al., *Phase II trial of trastuzumab in women with advanced or recurrent, HER2-positive endometrial carcinoma: a Gynecologic Oncology Group study.* Gynecol Oncol, 2010. **116**(1): p. 15-20.
159. Bregar, A.J. and W.B. Growdon, *Emerging strategies for targeting PI3K in gynecologic cancer.* Gynecol Oncol, 2016. **140**(2): p. 333-44.
160. Reinbolt, R.E. and J.L. Hays, *The Role of PARP Inhibitors in the Treatment of Gynecologic Malignancies.* Front Oncol, 2013. **3**: p. 237.
161. Gargiulo, P., et al., *Tumor genotype and immune microenvironment in POLE-ultramutated and MSI-hypermutated Endometrial Cancers: New candidates for checkpoint blockade immunotherapy?* Cancer Treat Rev, 2016. **48**: p. 61-8.
162. Moreno-Bueno, G., et al., *Differential gene expression profile in endometrioid and nonendometrioid endometrial carcinoma: STK15 is frequently overexpressed and amplified in nonendometrioid carcinomas.* Cancer Res, 2003. **63**(18): p. 5697-702.
163. Cao, Q.J., et al., *Distinctive gene expression profiles by cDNA microarrays in endometrioid and serous carcinomas of the endometrium.* Int J Gynecol Pathol, 2004. **23**(4): p. 321-9.
164. Fles, R., et al., *Genomic profile of endometrial tumors depends on morphological subtype, not on tamoxifen exposure.* Genes Chromosomes Cancer, 2010. **49**(8): p. 699-710.
165. Esteller, M., et al., *hMLH1 promoter hypermethylation is an early event in human endometrial tumorigenesis.* Am J Pathol, 1999. **155**(5): p. 1767-72.
166. Salvesen, H.B., et al., *Methylation of hMLH1 in a population-based series of endometrial carcinomas.* Clin Cancer Res, 2000. **6**(9): p. 3607-13.

167. Basil, J.B., et al., *Clinical significance of microsatellite instability in endometrial carcinoma.* Cancer, 2000. **89**(8): p. 1758-64.
168. Risinger, J.I., et al., *PTEN mutation in endometrial cancers is associated with favorable clinical and pathologic characteristics.* Clin Cancer Res, 1998. **4**(12): p. 3005-10.
169. Nagase, S., et al., *Deletion mapping on chromosome 10q25-q26 in human endometrial cancer.* Br J Cancer, 1996. **74**(12): p. 1979-83.
170. Oda, K., et al., *High frequency of coexistent mutations of PIK3CA and PTEN genes in endometrial carcinoma.* Cancer Res, 2005. **65**(23): p. 10669-73.
171. Salvesen, H.B., et al., *Integrated genomic profiling of endometrial carcinoma associates aggressive tumors with indicators of PI3 kinase activation.* Proc Natl Acad Sci U S A, 2009. **106**(12): p. 4834-9.
172. Urick, M.E., et al., *PIK3R1 (p85alpha) is somatically mutated at high frequency in primary endometrial cancer.* Cancer Res, 2011. **71**(12): p. 4061-7.
173. Lax, S.F., et al., *The frequency of p53, K-ras mutations, and microsatellite instability differs in uterine endometrioid and serous carcinoma: evidence of distinct molecular genetic pathways.* Cancer, 2000. **88**(4): p. 814-24.
174. Schlosshauer, P.W., L.H. Ellenson, and R.A. Soslow, *Beta-catenin and E-cadherin expression patterns in high-grade endometrial carcinoma are associated with histological subtype.* Mod Pathol, 2002. **15**(10): p. 1032-7.
175. Stefansson, I.M., H.B. Salvesen, and L.A. Akslen, *Prognostic impact of alterations in P-cadherin expression and related cell adhesion markers in endometrial cancer.* J Clin Oncol, 2004. **22**(7): p. 1242-52.
176. Guan, B., et al., *Mutation and loss of expression of ARID1A in uterine low-grade endometrioid carcinoma.* Am J Surg Pathol, 2011. **35**(5): p. 625-32.
177. Wiegand, K.C., et al., *Loss of BAF250a (ARID1A) is frequent in high-grade endometrial carcinomas.* J Pathol, 2011. **224**(3): p. 328-33.
178. Nagendra, D.C., et al., *PPP2R1A mutations are common in the serous type of endometrial cancer.* Mol Carcinog, 2012. **51**(10): p. 826-31.
179. Morrison, C., et al., *HER-2 is an independent prognostic factor in endometrial cancer: association with outcome in a large cohort of surgically staged patients.* J Clin Oncol, 2006. **24**(15): p. 2376-85.
180. Engelsen, I.B., et al., *Pathologic expression of p53 or p16 in preoperative curettage specimens identifies high-risk endometrial carcinomas.* Am J Obstet Gynecol, 2006. **195**(4): p. 979-86.
181. Salvesen, H.B., S. Das, and L.A. Akslen, *Loss of nuclear p16 protein expression is not associated with promoter methylation but defines a subgroup of aggressive endometrial carcinomas with poor prognosis.* Clin Cancer Res, 2000. **6**(1): p. 153-9.
182. Moreno-Bueno, G., et al., *Abnormalities of E- and P-cadherin and catenin (beta-, gamma-catenin, and p120ctn) expression in endometrial cancer and endometrial atypical hyperplasia.* J Pathol, 2003. **199**(4): p. 471-8.
183. Salvesen, H.B., I.S. Haldorsen, and J. Trovik, *Markers for individualised therapy in endometrial carcinoma.* Lancet Oncol, 2012. **13**(8): p. e353-61.
184. Marchio, C., et al., *PIKing the type and pattern of PI3K pathway mutations in endometrioid endometrial carcinomas.* Gynecol Oncol, 2015. **137**(2): p. 321-8.
185. Catasus, L., et al., *PIK3CA mutations in the kinase domain (exon 20) of uterine endometrial adenocarcinomas are associated with adverse prognostic parameters.* Mod Pathol, 2008. **21**(2): p. 131-9.
186. Sasaki, H., et al., *Mutation of the Ki-ras protooncogene in human endometrial hyperplasia and carcinoma.* Cancer Res, 1993. **53**(8): p. 1906-10.
187. Doll, A., et al., *Novel molecular profiles of endometrial cancer-new light through old windows.* J Steroid Biochem Mol Biol, 2008. **108**(3-5): p. 221-9.

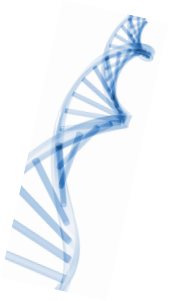
188. Liang, H., et al., *Whole-exome sequencing combined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer.* Genome Res, 2012. **22**(11): p. 2120-9.
189. Mao, T.L., et al., *Loss of ARID1A expression correlates with stages of tumor progression in uterine endometrioid carcinoma.* Am J Surg Pathol, 2013. **37**(9): p. 1342-8.
190. Kuhn, E., et al., *Identification of molecular pathway aberrations in uterine serous carcinoma by genome-wide analyses.* J Natl Cancer Inst, 2012. **104**(19): p. 1503-13.
191. Le Gallo, M., et al., *Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes.* Nat Genet, 2012. **44**(12): p. 1310-5.
192. Zhao, S., et al., *Landscape of somatic single-nucleotide and copy-number mutations in uterine serous carcinoma.* Proc Natl Acad Sci U S A, 2013. **110**(8): p. 2916-21.
193. Garcia-Sanz, P., et al., *Chromatin remodelling and DNA repair genes are frequently mutated in endometrioid endometrial carcinoma.* Int J Cancer, 2016.
194. Choi, Y.J., et al., *Genomic landscape of endometrial stromal sarcoma of uterus.* Oncotarget, 2015. **6**(32): p. 33319-28.
195. Rosa-Rosa, J.M., et al., *Molecular genetic heterogeneity in undifferentiated endometrial carcinomas.* Mod Pathol, 2016.
196. Cancer Genome Atlas Research, N., et al., *Integrated genomic characterization of endometrial carcinoma.* Nature, 2013. **497**(7447): p. 67-73.
197. Choi, Y., et al., *Predicting the functional effect of amino acid substitutions and indels.* PLoS One, 2012. **7**(10): p. e46688.
198. Sim, N.L., et al., *SIFT web server: predicting effects of amino acid substitutions on proteins.* Nucleic Acids Res, 2012. **40**(Web Server issue): p. W452-7.
199. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.* Brief Bioinform, 2013. **14**(2): p. 178-92.
200. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
201. Rios, D., et al., *A database and API for variation, dense genotyping and resequencing data.* BMC Bioinformatics, 2010. **11**: p. 238.
202. Gonzalez-Perez, A. and N. Lopez-Bigas, *Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel.* Am J Hum Genet, 2011. **88**(4): p. 440-9.
203. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
204. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
205. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.* Nucleic Acids Res, 2009. **37**(1): p. 1-13.
206. Smoot, M.E., et al., *Cytoscape 2.8: new features for data integration and network visualization.* Bioinformatics, 2011. **27**(3): p. 431-2.
207. Weinreb, I., et al., *Hotspot activating PRKD1 somatic mutations in polymorphous low-grade adenocarcinomas of the salivary glands.* Nat Genet, 2014. **46**(11): p. 1166-9.
208. Guerini-Rocco, E., et al., *The repertoire of somatic genetic alterations of acinic cell carcinomas of the breast: an exploratory, hypothesis-generating study.* J Pathol, 2015. **237**(2): p. 166-78.
209. Schultheis, A.M., et al., *Massively Parallel Sequencing-Based Clonality Analysis of Synchronous Endometrioid Endometrial and Ovarian Carcinomas.* J Natl Cancer Inst, 2016. **108**(6): p. djv427.

210. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
211. Cibulskis, K., et al., *Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.* Nat Biotechnol, 2013. **31**(3): p. 213-9.
212. Saunders, C.T., et al., *Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs.* Bioinformatics, 2012. **28**(14): p. 1811-7.
213. Koboldt, D.C., et al., *VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing.* Genome Res, 2012. **22**(3): p. 568-76.
214. Ha, G., et al., *TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data.* Genome Res, 2014. **24**(11): p. 1881-93.
215. Carter, S.L., et al., *Absolute quantification of somatic DNA alterations in human cancer.* Nat Biotechnol, 2012. **30**(5): p. 413-21.
216. Martincorena, I. and P.J. Campbell, *Somatic mutation in cancer and normal cells.* Science, 2015. **349**(6255): p. 1483-9.
217. Nik-Zainal, S., et al., *Mutational processes molding the genomes of 21 breast cancers.* Cell, 2012. **149**(5): p. 979-93.
218. Helleday, T., S. Eshtad, and S. Nik-Zainal, *Mechanisms underlying mutational signatures in human cancers.* Nat Rev Genet, 2014. **15**(9): p. 585-98.
219. Marquard, A.M., et al., *TumorTracer: a method to identify the tissue of origin from the somatic mutations of a tumor specimen.* BMC Med Genomics, 2015. **8**: p. 58.
220. Grondin, C.J., et al., *Advancing Exposure Science through Chemical Data Curation and Integration in the Comparative Toxicogenomics Database.* Environ Health Perspect, 2016. **124**(10): p. 1592-1599.
221. Davis, A.P., et al., *The Comparative Toxicogenomics Database: update 2017.* Nucleic Acids Res, 2017. **45**(D1): p. D972-D978.
222. Kuhn, M., et al., *STITCH: interaction networks of chemicals and proteins.* Nucleic Acids Res, 2008. **36**(Database issue): p. D684-8.
223. Schliep, K.P., *phangorn: phylogenetic analysis in R.* Bioinformatics, 2011. **27**(4): p. 592-3.
224. Buza, N., J. Ziai, and P. Hui, *Mismatch repair deficiency testing in clinical practice.* Expert Rev Mol Diagn, 2016. **16**(5): p. 591-604.
225. Boland, C.R., et al., *A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer.* Cancer Res, 1998. **58**(22): p. 5248-57.
226. Kallioniemi, A., et al., *Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors.* Science, 1992. **258**(5083): p. 818-21.
227. Hong, Y.S., et al., *Bortezomib induces G2-M arrest in human colon cancer cells through ROS-inducible phosphorylation of ATM-CHK1.* Int J Oncol, 2012. **41**(1): p. 76-82.
228. Gonzalez-Martin, A., et al., *Efficacy and safety results from OCTAVIA, a single-arm phase II study evaluating front-line bevacizumab, carboplatin and weekly paclitaxel for ovarian cancer.* Eur J Cancer, 2013. **49**(18): p. 3831-8.
229. Sims, D., et al., *Sequencing depth and coverage: key considerations in genomic analyses.* Nat Rev Genet, 2014. **15**(2): p. 121-32.
230. Tritz, D., et al., *Loss of heterozygosity in usual and special variant carcinomas of the endometrium.* Hum Pathol, 1997. **28**(5): p. 607-12.
231. Chang, M.T., et al., *Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity.* Nat Biotechnol, 2016. **34**(2): p. 155-63.
232. Gilks, C.B., E. Oliva, and R.A. Soslow, *Poor interobserver reproducibility in the diagnosis of high-grade endometrial carcinoma.* Am J Surg Pathol, 2013. **37**(6): p. 874-81.

233. Santacana, M., et al., *A 9-protein biomarker molecular signature for predicting histologic type in endometrial carcinoma by immunohistochemistry.* Hum Pathol, 2014. **45**(12): p. 2394-403.
234. Moreno-Bueno, G., et al., *Molecular alterations associated with cyclin D1 overexpression in endometrial cancer.* Int J Cancer, 2004. **110**(2): p. 194-200.
235. Musgrove, E.A., et al., *Cyclin D as a therapeutic target in cancer.* Nat Rev Cancer, 2011. **11**(8): p. 558-72.
236. Navin, N.E. and J. Hicks, *Tracing the tumor lineage.* Mol Oncol, 2010. **4**(3): p. 267-83.
237. Aparicio, S. and E. Mardis, *Tumor heterogeneity: next-generation sequencing enhances the view from the pathologist's microscope.* Genome Biol, 2014. **15**(9): p. 463.
238. Ryu, D., et al., *Deciphering intratumor heterogeneity using cancer genome analysis.* Hum Genet, 2016. **135**(6): p. 635-42.
239. Hiley, C., et al., *Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine.* Genome Biol, 2014. **15**(8): p. 453.
240. Roth, A., et al., *PyClone: statistical inference of clonal population structure in cancer.* Nat Methods, 2014. **11**(4): p. 396-8.
241. Qiao, Y., et al., *SubcloneSeeker: a computational framework for reconstructing tumor clone structure for cancer variant interpretation and prioritization.* Genome Biol, 2014. **15**(8): p. 443.
242. Navin, N.E., *Cancer genomics: one cell at a time.* Genome Biol, 2014. **15**(8): p. 452.
243. Khoo, B.L., et al., *Single-cell profiling approaches to probing tumor heterogeneity.* Int J Cancer, 2016. **139**(2): p. 243-55.
244. Turke, A.B., et al., *Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC.* Cancer Cell, 2010. **17**(1): p. 77-88.
245. Diaz, L.A., Jr., et al., *The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers.* Nature, 2012. **486**(7404): p. 537-40.
246. Burrell, R.A., et al., *The causes and consequences of genetic heterogeneity in cancer evolution.* Nature, 2013. **501**(7467): p. 338-45.
247. Gordon, D.J., B. Resio, and D. Pellman, *Causes and consequences of aneuploidy in cancer.* Nat Rev Genet, 2012. **13**(3): p. 189-203.
248. Freed-Pastor, W.A. and C. Prives, *Mutant p53: one name, many proteins.* Genes Dev, 2012. **26**(12): p. 1268-86.
249. Compton, D.A., *Mechanisms of aneuploidy.* Curr Opin Cell Biol, 2011. **23**(1): p. 109-13.
250. Beloribi-Djefaflia, S., S. Vasseur, and F. Guillaumond, *Lipid metabolic reprogramming in cancer cells.* Oncogenesis, 2016. **5**: p. e189.
251. Soon, P.S., et al., *Loss of heterozygosity of 17p13, with possible involvement of ACADVL and ALOX15B, in the pathogenesis of adrenocortical tumors.* Ann Surg, 2008. **247**(1): p. 157-64.
252. Levesque, E., et al., *Steroidogenic germline polymorphism predictors of prostate cancer progression in the estradiol pathway.* Clin Cancer Res, 2014. **20**(11): p. 2971-83.
253. Seo, Y.K., et al., *SULT2B1b sulfotransferase: induction by vitamin D receptor and reduced expression in prostate cancer.* Mol Endocrinol, 2013. **27**(6): p. 925-39.
254. Zhu, L. and M. Bakovic, *Breast cancer cells adapt to metabolic stress by increasing ethanolamine phospholipid synthesis and CTP:ethanolaminephosphate cytidylyltransferase-Pcyt2 activity.* Biochem Cell Biol, 2012. **90**(2): p. 188-99.
255. Gizard, F., et al., *TReP-132 is a novel progesterone receptor coactivator required for the inhibition of breast cancer cell growth and enhancement of differentiation by progesterone.* Mol Cell Biol, 2006. **26**(20): p. 7632-44.
256. Goldstein, I. and V. Rotter, *Regulation of lipid metabolism by p53 - fighting two villains with one sword.* Trends Endocrinol Metab, 2012. **23**(11): p. 567-75.

257. Notta, F., et al., *A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns.* Nature, 2016. **538**(7625): p. 378-382.
258. Anglesio, M.S., et al., *Synchronous Endometrial and Ovarian Carcinomas: Evidence of Clonality.* J Natl Cancer Inst, 2016. **108**(6): p. djv428.
259. Rebhandl, S., et al., *AID/APOBEC deaminases and cancer.* Oncoscience, 2015. **2**(4): p. 320-33.
260. Kanu, N., et al., *DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer.* Genome Biol, 2016. **17**(1): p. 185.
261. Depreeuw, J., et al., *Characterization of patient-derived tumor xenograft models of endometrial cancer for preclinical evaluation of targeted therapies.* Gynecol Oncol, 2015. **139**(1): p. 118-26.
262. Garralda, E., et al., *Integrated next-generation sequencing and avatar mouse models for personalized cancer treatment.* Clin Cancer Res, 2014. **20**(9): p. 2476-84.
263. Cleary, A.S., et al., *Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers.* Nature, 2014. **508**(7494): p. 113-7.
264. Wang, D., et al., *Molecular heterogeneity of non-small cell lung carcinoma patient-derived xenografts closely reflect their primary tumors.* Int J Cancer, 2017. **140**(3): p. 662-673.
265. Hao, C., et al., *Gene mutations in primary tumors and corresponding patient-derived xenografts derived from non-small cell lung cancer.* Cancer Lett, 2015. **357**(1): p. 179-85.
266. Choi, Y.Y., et al., *Establishment and characterisation of patient-derived xenografts as paraclinical models for gastric cancer.* Sci Rep, 2016. **6**: p. 22172.
267. Tomasella, A., et al., *The isopeptidase inhibitor 2cPE triggers proteotoxic stress and ATM activation in chronic lymphocytic leukemia cells.* Oncotarget, 2016. **7**(29): p. 45429-45443.
268. Bang, Y.J., et al., *Randomized, Double-Blind Phase II Trial With Prospective Classification by ATM Protein Level to Evaluate the Efficacy and Tolerability of Olaparib Plus Paclitaxel in Patients With Recurrent or Metastatic Gastric Cancer.* J Clin Oncol, 2015. **33**(33): p. 3858-65.
269. Cremona, C.A. and A. Behrens, *ATM signalling and cancer.* Oncogene, 2014. **33**(26): p. 3351-60.
270. Hoang, B., et al., *SGK Kinase Activity in Multiple Myeloma Cells Protects against ER Stress Apoptosis via a SEK-Dependent Mechanism.* Mol Cancer Res, 2016. **14**(4): p. 397-407.
271. Hambley, B., P.F. Caimi, and B.M. William, *Bortezomib for the treatment of mantle cell lymphoma: an update.* Ther Adv Hematol, 2016. **7**(4): p. 196-208.
272. Scott, K., et al., *Bortezomib for the treatment of multiple myeloma.* Cochrane Database Syst Rev, 2016. **4**: p. CD010816.
273. Dolcet, X., et al., *Proteasome inhibitors induce death but activate NF-kappaB on endometrial carcinoma cell lines and primary culture explants.* J Biol Chem, 2006. **281**(31): p. 22118-30.
274. Llobet, D., et al., *Antioxidants block proteasome inhibitor function in endometrial carcinoma cells.* Anticancer Drugs, 2008. **19**(2): p. 115-24.
275. Shen, Y., et al., *Bortezomib induces apoptosis of endometrial cancer cells through microRNA-17-5p by targeting p21.* Cell Biol Int, 2013. **37**(10): p. 1114-21.
276. Sorolla, A., et al., *Blockade of NFkappaB activity by Sunitinib increases cell death in Bortezomib-treated endometrial carcinoma cells.* Mol Oncol, 2012. **6**(5): p. 530-41.
277. Yeramian, A., et al., *Inhibition of activated receptor tyrosine kinases by Sunitinib induces growth arrest and sensitizes melanoma cells to Bortezomib by blocking Akt pathway.* Int J Cancer, 2012. **130**(4): p. 967-78.
278. Bestvina, C.M. and G.F. Fleming, *Chemotherapy for Endometrial Cancer in Adjuvant and Advanced Disease Settings.* Oncologist, 2016. **21**(10): p. 1250-1259.

279. Sleijfer, S., J. Bogaerts, and L.L. Siu, *Designing transformative clinical trials in the cancer genome era.* J Clin Oncol, 2013. **31**(15): p. 1834-41.
280. Piccart-Gebhart, M.J., et al., *Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer.* N Engl J Med, 2005. **353**(16): p. 1659-72.
281. Rosell, R., et al., *Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial.* Lancet Oncol, 2012. **13**(3): p. 239-46.
282. Bartsch, R. and E. de Azambuja, *I-SPY 2: optimising cancer drug development in the 21st century.* ESMO Open, 2016. **1**(5): p. e000113.
283. Jamal-Hanjani, M., et al., *Tracking genomic cancer evolution for precision medicine: the lung TRACERx study.* PLoS Biol, 2014. **12**(7): p. e1001906.
284. Choi, Y.J., et al., *Intraindividual genomic heterogeneity of high-grade serous carcinoma of the ovary and clinical utility of ascitic cancer cells for mutation profiling.* J Pathol, 2017. **241**(1): p. 57-66.
285. Yi, X., et al., *The Feasibility of Using Mutation Detection in ctDNA to Assess Tumor Dynamics.* Int J Cancer, 2017.
286. Alonso-Alconada, L., et al., *Molecular profiling of circulating tumor cells links plasticity to the metastatic process in endometrial cancer.* Mol Cancer, 2014. **13**: p. 223.
287. Gultekin, M., et al., *Comparison of FIGO 1988 and 2009 staging systems for endometrial carcinoma.* Med Oncol, 2012. **29**(4): p. 2955-62.
288. Page, B.R., et al., *Does the FIGO 2009 endometrial cancer staging system more accurately correlate with clinical outcome in different histologies? Revised staging, endometrial cancer, histology.* Int J Gynecol Cancer, 2012. **22**(4): p. 593-8.



ANNEX I: Publications

During this thesis, the following publications have been generated:

1. **Mota, A.**, Trivino, J. C., Rojo-Sebastian, A., Martinez-Ramirez, A., Chiva, L., Gonzalez-Martin, A., Garcia, J. F., Garcia-Sanz, P. & Moreno-Bueno, G. Intratumor heterogeneity in TP53 null High Grade Serous Ovarian Carcinoma progression. BMC Cancer. 2015 Nov 30;15:940. doi: 10.1186/s12885-015-1952-z.

2. **Mota, A.**, Colas, E., Garcia-Sanz, P., Campoy, I., Rojo-Sebastian, A., Gatius, S., Garcia, A., Chiva, L., Alonso, S., Gil-Moreno, A., Gonzalez-Tallada, X., Diaz-Feijoo, B., Vidal, A., Ziober-Malinowska, P., Bobinski, M., Lopez-Lopez, R., Abal, M., Reventos, J., Matias-Guiu, X. & Moreno-Bueno, G. Genetic analysis of uterine aspirates improves the diagnostic value and captures the intratumor heterogeneity of endometrial cancers. Mod Pathol. 2016 Sep 2. doi: 10.1038/modpathol.2016.143.

3. **Mota, A.**, Burke, KA., Selenic, P., Colas, E., Moiola, CP., Perez-Lopez, M., Triviño, JC., Rojo-Sebastian, A., Santacana, M., Ortega, E., Gil-Moreno, A., Reis-Filho, JS., Weilget, B., Matias-Guiu, X., Moreno-Bueno, G. Decoding the genomic profile of an ambiguous endometrial carcinoma and its derived preclinical models (In preparation).

4. **Mota, A.,** Burke, KA., Selenic, P., Garcia-Sanz, P., Colas, E., Gatius, S., Garcia, A., Rojo-Sebastian, A., Vidal, A., Gil-Moreno, A., Reis-Filho, JS., Matias-Guiu, X., Weilget, B., Moreno-Bueno, G. Intratumor genomic heterogeneity landscape of metastatic endometrial carcinomas unravels different clonal evolution (In preparation).

In addition, the doctoral student has participated in the following collaborative publications:

5. Santacana, M., Maiques, O., Valls, J., Gatius, S., Abo, A. I., Lopez-Garcia, M. A., **Mota, A.**, Reventos, J., Moreno-Bueno, G., Palacios, J., Bartosch, C., Dolcet, X. & Matias-Guiu, X. A 9-protein biomarker molecular signature for predicting histologic type in endometrial carcinoma by immunohistochemistry. Hum Pathol. 2014 Dec;45(12):2394-403. doi: 10.1016/j.humpath.2014.06.031.

6. Hergueta-Redondo, M., Sarrio, D., Molina-Crespo, A., Megias, D., **Mota, A.**, Rojo-Sebastian, A., Garcia-Sanz, P., Morales, S., Abril, S., Cano, A., Peinado, H. & Moreno-Bueno, G. Gasdermin-b promotes invasion and metastasis in breast cancer cells. PLoS One. 2014 Mar 27;9(3):e90099. doi: 10.1371/journal.pone.0090099.

7. Romero-Perez, L., Garcia-Sanz, P., **Mota, A.**, Leskela, S., Hergueta-Redondo, M., Diaz-Martin, J., Lopez-Garcia, M. A., Castilla, M. A., Martinez-Ramirez, A., Soslow, R. A., Matias-Guiu, X., Moreno-Bueno, G. & Palacios, J. A role for the transducer of the hippo pathway, taz, in the development of aggressive types of endometrial cancer. Mod Pathol. 2015 Nov;28(11):1492-503. doi: 10.1038/modpathol.2015.102.

8. Hellner, K., Miranda, F., Fotso Chedom, D., Herrero-Gonzalez, S., Hayden, DM., Tearle, R., Artibani M., KaramiNejadRanjbar, M., Williams, R., Gaitskell, K., Elorbany, S., Xu, R., Laios, A., Buiga, P., Ahmed, K., Dhar, S., Zhang, RY., Campo, L., Myers, KA., Lozano, M., Ruiz-Miró, M., Gatius, S., **Mota, A.**, Moreno-Bueno, G., Matias-Guiu, X., Benítez, J., Witty, L., McVean, G., Leedham, S., Tomlinson, I., Drmanac, R., Cazier, JB., Klein, R., Dunne, K., Bast, RC Jr., Kennedy, SH., Hassan, B., Lise, S., Garcia, MJ., Peters, BA., Yau, C., Sauka-Spengler, T., Ahmed, AA. Premalignant SOX2 overexpression in the fallopian tubes of ovarian cancer patients: Discovery and validation studies. EBioMedicine. 2016 Aug;10:137-49. doi: 10.1016/j.ebiom.2016.06.048

9. Hergueta-Redondo, M., Sarrio, D., Molina-Crespo, Á., Vicario, R., Bernadó-Morales, C., Martínez, L., Rojo-Sebastián, A., Serra-Musach, J., **Mota, A**., Martínez-Ramírez, Á., Castilla, MÁ., González-Martin, A., Pernas, S., Cano, A., Cortes, J., Nuciforo, PG., Peg, V., Palacios, J., Pujana, MÁ., Arribas, J., Moreno-Bueno G. Gasdermin B expression predicts poor clinical outcome in HER2-positive breast cancer. Oncotarget. 2016 Jul 22. doi: 10.18632/oncotarget.10787.

10. Santacana, M., Coronado, P., Matias-Guiu, X., Romero, I., Casado, A., Gil-Moreno, A., Dosil, MA., **Mota, A**., Moreno-Bueno, G., Dolcet, X., Llombart-Cussac, A., Poveda, A. Biological Effects of Temsirolimus on the mTOR Pathway in Endometrial Carcinoma: A Pharmacodynamic Phase II Study. Int J Gynecol Cancer. 2016 Jun 2. doi: 10.1097/IGC.0000000000000715.

11. Piscuoglio, S., Geyer, FC., Burke, KA., Murray, M., Ng, CK., **Mota, A.,** Marchio, C., Berman, SH., Norton, L., Brogi, E., Weigelt, B., Reis-Filho, JS. Massively Parallel Sequencing Analysis of Synchronous Fibroepithelial Lesions Supports the Concept of Progression from Fibroadenoma to Phyllodes Tumor. NPJ Breast Cancer. 2016 Nov 16. doi:10.1038/npjbcancer.2016.35.

12. García-Sanz P, Triviño JC, **Mota A**, Pérez López M, Colás E, Rojo-Sebastián A, GArcía Á, Gatius S, Ruiz M, Prat J, López-López R, Abal M, Gil-Moreno A, Reventós J, Matias-Guiu X, Moreno-Bueno G. Chromatin remodelling and DNA repair genes are frequently mutated in endometrioid endometrial carcinoma. Int J Cancer. 2016 Dec 20. doi: 10.1002/ijc.30573.

Mota *et al. BMC Cancer* (2015) 15:940
DOI 10.1186/s12885-015-1952-z

BMC Cancer

RESEARCH ARTICLE Open Access

Intra-tumor heterogeneity in TP53 null High Grade Serous Ovarian Carcinoma progression

Alba Mota[1,2], Juan Carlos Triviño[3], Alejandro Rojo-Sebastian[4], Ángel Martínez-Ramírez[5], Luis Chiva[6], Antonio González-Martín[7], Juan F. Garcia[2,4], Pablo Garcia-Sanz[1,2*] and Gema Moreno-Bueno[1,2*]

Abstract

Background: High grade serous ovarian cancer is characterised by high initial response to chemotherapy but poor outcome in the long term due to acquired resistance. One of the main genetic features of this disease is *TP53* mutation. The majority of *TP53* mutated tumors harbor missense mutations in this gene, correlated with p53 accumulation. *TP53* null tumors constitute a specific subgroup characterised by nonsense, frameshift or splice-site mutations associated to complete absence of p53 expression. Different studies show that this kind of tumors may have a worse prognosis than other *TP53* mutated HGSC.

Methods: In this study, we sought to characterise the intra-tumor heterogeneity of a *TP53* null HGSC consisting of six primary tumor samples, two intra-pelvic and four extra-pelvic recurrences using exome sequencing and comparative genome hybridisation.

Results: Significant heterogeneity was found among the different tumor samples, both at the mutational and copy number levels. Exome sequencing identified 102 variants, of which only 42 were common to all three samples; whereas 7 of the 18 copy number changes found by CGH analysis were presented in all samples. Sanger validation of 20 variants found by exome sequencing in additional regions of the primary tumor and the recurrence allowed us to establish a sequence of the tumor clonal evolution, identifying those populations that most likely gave rise to recurrences and genes potentially involved in this process, like *GPNMB* and *TFDP1*. Using functional annotation and network analysis, we identified those biological functions most significantly altered in this tumor. Remarkably, unexpected functions such as microtubule-based movement and lipid metabolism emerged as important for tumor development and progression, suggesting its potential interest as therapeutic targets.

Conclusions: Altogether, our results shed light on the clonal evolution of the distinct tumor regions identifying the most aggressive subpopulations and at least some of the genes that may be implicated in its progression and recurrence, and highlights the importance of considering intra-tumor heterogeneity when carrying out genetic and genomic studies, especially when these are aimed to diagnostic procedures or to uncover possible therapeutic strategies.

Keywords: Ovary cancer, High grade serous carcinoma, *TP53* null, Intra-tumor heterogeneity

* Correspondence: pgarcia@mdanderson.es; gmoreno@iib.uam.es
[1]Departamento de Bioquímica, Universidad Autónoma de Madrid (UAM), Instituto de Investigaciones Biomédicas "Alberto Sols" (CSIC-UAM), IdiPAZ, Madrid, Spain
Full list of author information is available at the end of the article

Background

Ovarian cancer is the most common cause of death from gynecological malignancies. Around seventy percent of ovarian cancers are histologically classified as high-grade serous carcinoma (HGSC). The standard treatment for these tumors is cytoreductive surgery followed by platinum-taxane chemotherapy. Although the initial response rate is higher than 80 %, the majority of patients relapses within five years and die due to chemo-resistant disease [1, 2].

HGSCs are characterised by nearly universal *TP53* gene mutation, present in more than 95 % of the cases [3]. The most common *TP53* abnormalities are missense mutations, which induce nuclear accumulation of the mutant protein with a strongly positive IHC staining. However, approximately 30 % of somatic *TP53* mutations are considered null mutations that lead to complete absence of p53 protein due to nonsense, frameshift or splicing junction mutations [4]. The prognosis value of *TP53* status is a controversial issue [5]. Nevertheless, several studies support that tumors with *TP53* null mutations present a worse outcome compared with those in which *TP53* harbors mutations involving overexpression. This would happen not only in ovarian cancers [6–9], but also in breast, colorectal and head and neck cancers and leukemia [10].

The Cancer Genome Atlas (TCGA) consortium has enabled a deeper understanding of HGSCs using an integrated genomic approach for the analysis of 316 tumors. This study has revealed that, with the exception of *TP53* mutation, present in 96 % of the tumors, recurrent mutations are not common in HGSCs. Nonetheless the hallmark of HGSCs was numerous somatic copy number alterations, with more than 100 recurrent amplifications and deletions identified [11]. However, this analysis only considered primary tumors samples, regardless of subsequent recurrences. In fact, there are few studies that have analysed from a genomic point of view the evolution of these tumors using paired samples of primary tumor and post-treatment relapse. According to some of these studies, primary and recurrent diseases would be genetically similar [12–14]. Conversely, other studies have shown a high degree of intra-tumor heterogeneity, which would allow clonal evolution and the successive tumor progression and resistance to chemotherapy [15, 16]. In fact, intra-tumor heterogeneity has been shown to be intrinsic to primary tumor and not just a result of chemotherapy treatment [14, 17].

To the best of our knowledge, none of these studies considered *TP53* null tumors as an individual clinical entity. In the present study, whole-exome sequencing and comparative genomic hybridisation was performed on primary tumor and recurrence implants of a *TP53* null patient, revealing significant heterogeneity between all samples. Selected variants detected in the genomic analysis were subsequently validated in multiple tumoral implants from the primary tumor and the recurrence. These data revealed intra-tumor heterogeneity in the primary tumor, which was reflected in the different recurrence metastases, suggesting a model of clonal evolution.

Methods

Samples selection

Tumoral samples from primary and recurrence disease were selected from a TP53 null case of HGSC. Biospecimens and clinical data were collected after written approval of the study by the Ethics Committee of the MD Anderson Cancer Center. A specific informed consent was obtained from the patient managed by the MD Anderson Foundation Biobank (record number B.0000745, ISCIII National Biobank Record) and included the authorization for data publication of individual clinical data and any accompanying images. A total of twelve samples were selected, three frozen and nine Formalin-Fixed Paraffin-Embedded (FFPE) tissue. Six primary tumor (P1-P6), two intra-pelvic recurrence (IR1-IR2) and four extra-pelvic recurrence (ER1-ER4) implants were analysed (Fig. 1a). All samples selected contained at least 80 % of tumoral cells. Non-tumoral mesothelium (N) was used as control tissue.

Immunohistochemistry

Immunostaining was performed on 2 μm FFPE tissue sections. Deparaffinisation and antigen retrieval were done in a PT Link module (Dako, Denmark) and immunohistochemistry was done in an Autostainer (Dako). Primary antibodies used were WT1 (clone 6 F-H2, RTU, 1:1, Dako), PTEN (clone 6H2.1, 1:100, Dako), p53 (clone DO-7, RTU, 1:1, Dako) and Ki-67 (clone MIB-1, RTU, 1:1, Dako). Amplification and visualisation of immune complexes and counterstaining were also performed in the Autostainer with the use of an EnVision FLEX+, Mouse (Dako, CA, USA). Slides were counterstained with hematoxylin and evaluated by two independent pathologists.

Whole-exome sequencing (WES)

Whole-exome sequencing was performed in four frozen samples (N, P1, IR1 and ER1) using the latest version of SureSelectXT Human All Exon V4 + UTR (71 Mb) enrichment kit and SOLiD high-throughput sequencing platform, according to manufacturer's protocols. Sequencing reactions were carried out to obtain read pairs (75 + 35 nt, paired-end). Data quality was estimated using parameters from SETS (SOLiD Exerimental Tracking System) software. For the bioinformatics analysis, see Additional file 1: Supplementary Materials.

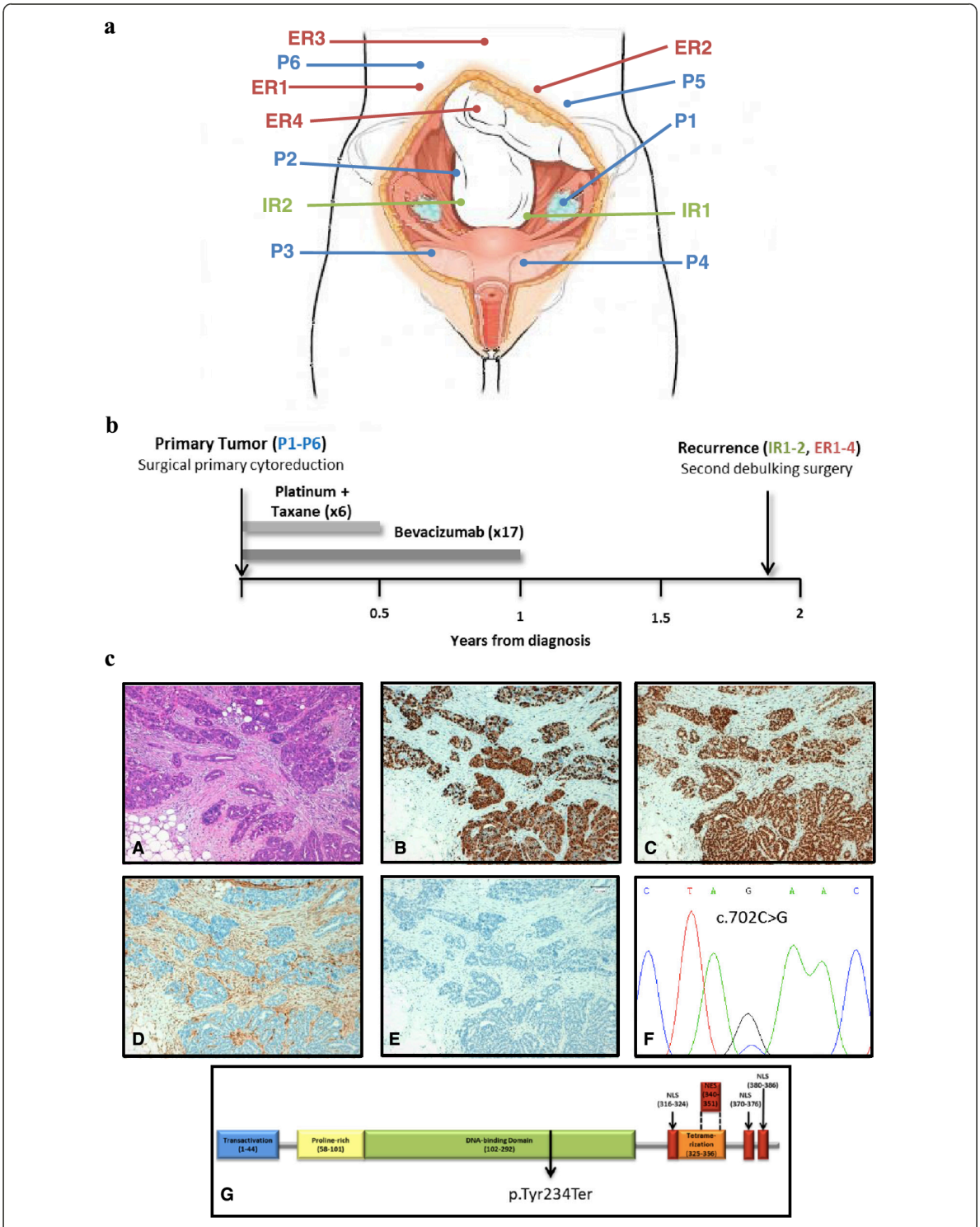


Fig. 1 (See legend on next page.)

(See figure on previous page.)
Fig. 1 Clinical case and sample description of an ovarian TP53 null HGSC. **a** Anatomical location of the primary tumor (blue) and recurrence samples (intra-pelvic samples in green and extra-pelvic samples in red). **b** Patient's clinical course. Grey squares indicate periods of treatment between primary tumor diagnosis and relapse. **c** Representative images of hematoxilin/eosin staining (A) and immunohistochemistries for WT1 (B), Ki-67 (C), PTEN (D) and p53 (E). TP53 mutation detected by Sanger sequencing (F) and its consequence in p53 protein (G)

Variant calling for Single Nucleotide Variants (SNVs) identification was performed using three different algorithms: VarScan2 [18], Mutect [19] and Bioscope. 'In-house' scripts were developed to combine the variants and filter possible technical artifacts. SNVs and Indel identified were annotated using the 'Applicaton Programming Interfaces' (APIs) from Ensembl v64 and several 'custom' scripts. SNVs and Indel with a minimum coverage of 10 and frequency above 0.1 percent were selected. Variants consequences were analysed using Condel [20] (http://bg.upf.edu/fannsdb/), SIFT [21] (http://sift.jcvi.org) and PolyPhen [22] (http://genetics.bwh.harvard.edu/pph2/) predictors.

SNVs Validation

A total of 25 variants detected by whole-exome sequencing were reanalysed by Sanger sequencing. PCR conditions and length of amplicons are indicated in Additional file 1: Table S2.

Comparative genomic hybridisation (CGH)

CGH was performed according to Kallioniemi et al. [23], with some modifications. Tumor and normal DNA were labeled using a nick translation kit (Abbott Molecular Inc.). In short, 200 ng of each labeled DNA was hybridised to normal female metaphase cells in the presence of 20–35 mg of Cot-1 DNA for 3 days. After washes, the chromosomes were counterstained with DAPI in an antifade solution. Analysis was performed with a Leica DM4500 epifluorescence microscope equipped with a CCD camera. A minimum of 15 metaphase cells per hybridisation per case were analysed by use of the CytoVision System with version 7.3.1 high-resolution CGH analysis software (Leica Biosystems, UK). The CGH profiles were compared to a dynamic standard reference interval on the basis of an average of normal cases, as previously described [24].

Fluorescence in situ hybridisation (FISH)

Fluorescence in situ hybridisation was performed as previously described [24]. Briefly, metaphase chromosomes were prepared directly from the same samples used for CGH. Slides were placed at 90 °C for 10 min, dehydrated through a series of ethanol washes, and denatured in the presence of a probe on a plate at 75 °C for 1 min. For detection of gene amplification, gene-specific probes for *PML* and *RARA* (used as control) (Vysis, Downers Grove, IL) were used. At least 100 interphase nuclei were analysed.

Functional annotation

The selected variants were annotated using David protocols [25, 26]. Briefly, this method allows the functional annotation of genes using different biological database as Biocarta (http://www.biocarta.com/), Gene Ontology (http://geneontology.org/), KEGG (http://www.genome.jp/kegg/pathway.html) and Reactome (http://www.reactome.org/). The interesting functional categories were selected using a p-Value threshold of 0.2. Finally, the correlation between genes and functional categories was represented in form of network using Cytoscape tool [27].

Hierarchical clustering

A hierarchical clustering method was applied to group the samples on the basis of similarities in mutation pattern. The unsupervised analyses were carried out using the SPSS 17.0 for statistical program (SPSS Inc., Chicago, IL) assuming Euclidean distances between mutations.

Results

Clinical history and case description

The patient of this study was diagnosed in 2009 at age 50 with a stage IIIC high grade serous carcinoma (HGSC) with extensive intra- and extra-pelvic peritoneal carcinomatosis. After surgical primary cytoreduction the patient was left with residual disease of less than 1 cm. Afterwards she was included in the OCTAVIA clinical trial [28] and treated with six cycles of weekly paclitaxel and carboplatin, with bevacizumab every 3 weeks for a total period of 12 months. Patient relapsed after 23 months of platinum free interval with numerous tumor implants in intra- and extra-pelvic peritoneum (Fig. 1a). The patient underwent a second debulking surgery without residual macroscopic disease followed by 6 cycles of carboplatin and pegylated liposomal doxorubicin hydrochloride (Fig. 1b).

Both primary and relapsed tumors showed papillary patterns with frequent necrosis, nuclear expression of WT1 and very high proliferative rate (90 % as determined by Ki-67 staining). Immunohistochemistry also showed lack of PTEN expression (suggestive of mutation) and complete absence of p53 staining (indicative of null mutation). Sanger sequencing showed a mutation in exon 7 of *TP53* resulting in a premature stop codon in the position 234 of the protein (Fig. 1c). According to these data, this tumor is classified as a *TP53* null high grade serous carcinoma.

Identification of somatic nucleotide variants by whole-exome sequencing shows differences in the mutational patterns of distinct tumor regions

In order to detect somatic nucleotide variants, whole-exome sequencing was performed in the primary tumor (P1), pararectal recurrence implant (IR1), ileal recurrence implant (ER1) and mesothelium as normal tissue (N) from a *TP53* null patient. After comparison of tumoral samples with the normal tissue, the bioinformatic analysis identified 102 variants, 99 SNVs and three deletions (Additional file 1: Table S1). Only 42 variants were common to all samples, comprising just 41 % of the total. Sample specific variants were identified in the three regions, as well as variants common to two samples (Fig. 2a). Notably, the range of frequencies found in the primary tumor variants suggested the existence of intra-tumor heterogeneity from the onset.

The consequence of each variant in the corresponding protein was analysed by the computational predictors Condel, SIFT and Polyphen [20–22]. A total of 80 variants presented a possibly damaging or deleterious consequence as predicted by at least one of them (Fig. 2b). Interestingly, the scattering of these variants reflects the unequal distribution of the mutations potentially implicated in tumor progression.

A selection of 25 variants detected by WES was reanalysed by Sanger sequencing. Four of them were not validated, showing a false positive rate of 16 % in the WES study (Additional file 1: Table S2). The *TP53* nonsense mutation was detected ubiquitously in the three samples analysed, consistent with the founder role of this gene in this type of cancer.

Copy Number Variations (CNVs) are unequally distributed among the different tumor regions

HGSCs are characterised by high chromosomal instability and widespread copy number changes. In order to analyse copy number variations (CNVs), Comparative Genomic Hybridisation (CGH) was performed in the primary tumor (P1) and two recurrence implants (IR1 and ER1), using normal tissue (N) as a reference. We identified changes in 18 chromosomal regions, with gain of material more frequent than loss (61 % vs. 39 %) (Fig. 3, Additional file 1: Table S3). A total of 7 changes (38 % of total) were common to the three samples, counting five gains (2q32q33, 3q22q29, 7q22q32, 8q12q24 and 11q14q22) and two diminished regions (6q25q27 and depletion of whole chromosome 4). P1 and ER1 shared 5 regions (gain of 1p22p35 and 9q31, and loss of 17p13 and depletion of whole chromosomes 12 and 16), while specific changes of each sample were detected (8p22p23 and 15q22q26 in P1; 16q24 in IR1; and 5p15, 10q22 and 22q13 in ER1). Interestingly, no changes exclusively common to P1 and IR1 or to IR1 and ER1 were found.

The analysis of additional tumor regions shows the existence of mutationally heterogeneous tumor clones

To further investigate intra-tumor heterogeneity, a subset of 20 of the variants detected by exome sequencing was validated by Sanger sequencing in the four original samples (N, P1, IR1, ER1) and in nine additional FFPE samples of tumor implants representing different regions of the primary tumor and recurrences (P2-P6, IR2, ER2-ER4) (Fig. 4a). For the location of the different tumor regions, see Fig. 1a. Variants analysed were selected to represent the different distributions identified in whole-exome sequencing: nine ubiquitous (*TP53*, *CSMD3*, *CTC1*, *FAP*, *KIF21B*, *LAMA2*, *SMG7*, *UBR3* and *ZFAT*), one shared by P1 and IR1 (*FRMPD1*), four shared by IR1 and ER1 (*HEPHL1*, *KIF21A*, *OR56B1* and *ZNF664*), three P1 specific (*CNOT1*, *PLXNA1* and *TRERF1*), two IR1 specific (*ROBO2* and *TFDP1*) and one ER1 specific (*GPNMB*). All these variants were potentially damaging, nonsense, frameshift or splice site variants (Additional file 1: Table S1).

As expected some variants were detected in all of the analysed samples, consistent with the exome sequencing results in which they were detected ubiquitously in P1, IR1 and ER1 samples (Fig. 4a). However, other variants, also ubiquitous according to exome sequencing, were not present in all the samples. Moreover, the *FRMPD1* variant, originally presented in primary tumor and IR1 samples and not in ER1, was current not only in all primary and intra-pelvic regions but also in the other extra-pelvic recurrence samples analysed.

On the other hand, some of the variants initially detected variants as characteristic of them were detected not only in new intra- and extra-pelvic recurrences, but also in primary tumor regions different from the one originally analysed. Moreover, *ZNF664* mutation, (common in IR1 and ER1 samples) was also found mutated in all of the recurrence samples but not in the primary tumor. Specific sample variants were also detected by Sanger sequencing, both for the primary tumor and for intra- and extra-pelvic recurrences, showing a certain degree of mutational heterogeneity among the different tumor locations.

Additionally, intra-tumor genomic heterogeneity was studied by fluorescence in situ hybridisation (FISH) in *PML* gene (15q22), exclusively altered in P1 sample regarding CGH results (Additional file 2: Figure S1). *PML* gain was observed by FISH in P1 but not in IR1 and ER1 samples, confirming CGH data. Furthermore, heterogeneous CNVs were observed in FFPE samples analysed. While P3, P4, P5 and ER1 samples showed 3 or 4 copies of *PML*; P2, P6, IR2 and ER4 presented between 5 and 8 copies. ER3 sample presented just one copy of *PML* per cell, indicating a deletion of this gene in this specific tumor sample (Additional file 2: Figure S1).

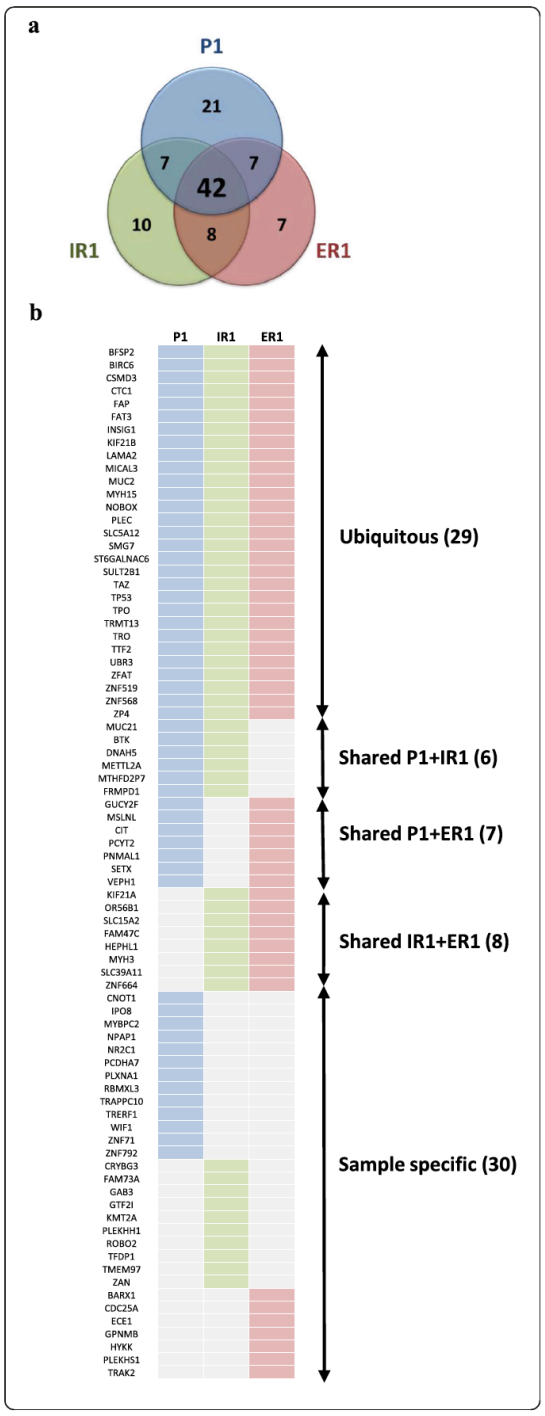


Fig. 2 Somatic variants detected by whole-exome sequencing in a TP53 null HGSC. **a** Venn diagram showing the distribution of the total number of variants among the different tumor samples (P1, primary tumor: blue; IR1, intrapelvic recurrence sample: green; ER1, extrapelvic recurrence sample: red). **b** Genes with missense, splice site, nonsense or frameshift mutations with a negative consequence in the corresponding protein and its distribution in the different analysed samples

Therefore intra-tumor heterogeneity was not only observed in genetic variations but also in genomic changes.

Hierarchical clustering according to mutation pattern allows determining clonal evolution

Since the mutational heterogeneity observed in this tumor is consistent with a clonal composition, we performed a hierarchical clustering analysis based in the distribution of the subset of 20 variants among the 12 different regions of the primary tumor and the two recurrences (Fig. 4b). This analysis shows the existence of two main branches, as shown in the diagram. Most of the recurrences regions would evolve from a primary tumor ancestor subpopulation that gives rise to the upper branch, originating P3, P4 and P5 regions of the primary tumor, the intra-pelvic recurrences and most of the extra-pelvic. Notably, ER2 seems to be more closely related to the intra-pelvic recurrence than to the other extra-pelvic regions. The lower branch contains P1 (the closest region to normal tissue), P2 and P6, and ER3 which would arise from this latter primary tumor region.

Remarkably, it is also possible to hierarchically cluster the three original tumor regions (P1, IR1 and ER1) according to the WES or CGH results. This tumor phylogenetic tree is equivalent to the one generated with the previous subset of 20 mutations, showing two main branches, one corresponding to the primary tumor (P1 region) and the other subdivided in two sub-branches which contain each of the recurrences samples (data not shown).

Functional annotation and network analysis identifies biological functions implicated in tumor development and progression

In order to analyse from a global point of view the biological and molecular functions implicated in the development and progression of this tumor, the functions of the genes mutated according to the exome sequencing were annotated using different biological databases as Biocarta, Gene Ontology, KEGG and Reactome (see Methods section). Network analysis using Cytoscape plugin [27] identified cell adhesion, cell cycle control, microtubule-related movement and transport, lipid metabolism and apoptosis as pathways containing multiple mutated genes (Fig. 5, Additional file 1: Table S4).

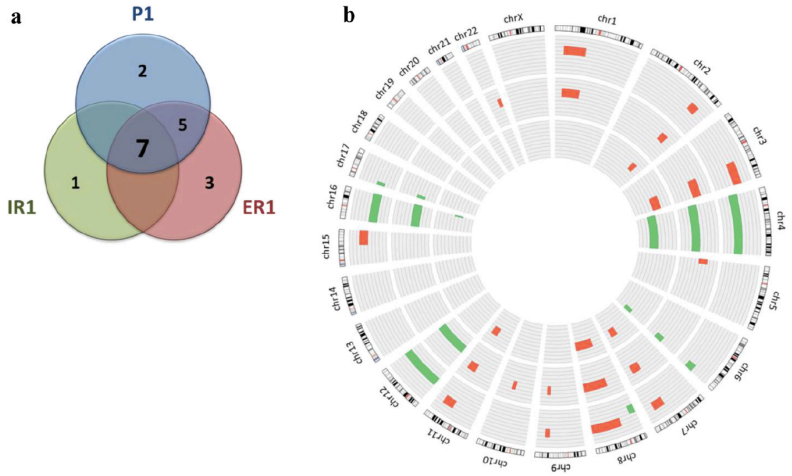


Fig. 3 Copy Number Variants (CNVs) detected by comparative genomic hybridisation (CGH) in a TP53 null HGSC. **a** Venn diagram showing the distribution of the total number of copy number changes among the different samples ((P1 primary tumor: blue; IR1,intrapelvic recurrence sample: green; ER1, extrapelvic recurrence sample: red). **b** Circes representation of the copy number changes. P1, ER1 and IR1 samples are displayed in concentric circles (from the outside to the inside). Enhanced regions are represented in red and diminished regions in green

Cell adhesion was consistently affected in all the analysed samples, both with mutations common to all samples and with shared or sample specific alterations. Something similar, but with a lower number of affected genes, occurred with cell cycle control, microtubule-based movement, cell division and chromosome partitioning and apoptosis. On the other hand, mutations in genes implicated in lipid metabolism were mainly ubiquitous, pointing to its possible involvement in tumor initiation. In addition, functional annotation was also performed in CGH data, identifying multiple genes implicated in the previous described pathways (data not shown).

Interestingly, although DNA repair pathway is frequently defective in these tumors, we did not identify mutated genes directly related to this pathway. The gene most closely related to this function was SETX, which participates in the defense against oxidative DNA damage [29]. However, CGH data showed that multiple mismatch repair genes, such as *RPA*, *POLE*, *POLD3* or *POLD4*, exhibited copy number changes in all the samples analysed (data not shown).

Discussion

Despite of initial response to treatment, the majority of HGSC patients will recur and die due to chemoresistant disease. It is well established that a defective status of the DNA repair machinery (especially *BRCA1/2* genes) correlates with a better response to platinum-based therapy [30]. Nevertheless, there is still a huge gap regarding the genetic and molecular factors that contribute to explain the high rate of therapy failure in this disease. Regarding the importance of *TP53* status in this disease, there are evidences pointing to a worse prognosis for tumors with null mutations compared with those in which overexpression inducing mutations are found, although this remains controversial [6–8, 10]. To this respect, a meta-analysis of the genomic data generated by the TCGA consortium regarding ovarian cancer shows decreased overall and progression-free survival for wild type *TP53* tumors with respect to mutant *TP53* cases [31]. Taking this into account, we divided the *TP53* mutant group analysed in the TCGA study into two subgroups: mutant and null tumors. *TP53* null tumors are characterised by complete absence of p53 protein, probably due to degradation of their mRNAs by nonsense-mediated RNA decay [32]. Accordingly, those cases with frameshift, splice site or nonsense mutations and with a mRNA level z-score lower than -1 were classified as *TP53* null tumors. This subgroup shows an intermediate overall survival between the wild-type and mutant groups, with nearly significant and equivalent differences between the different categories (data not shown). In view of that, a deeper understanding of *TP53* null tumors would be needed.

In this study we have analysed by WES and CGH the genetic and genomic alterations of an ovarian *TP53* null HGSC case consisting of the primary tumor, one intra-pelvic and another extra-pelvic recurrence. The exome sequencing analysis shows a significant mutational divergence between the primary tumor and both recurrences. Nevertheless, this degree of divergence decreases when other tumor regions are screened by

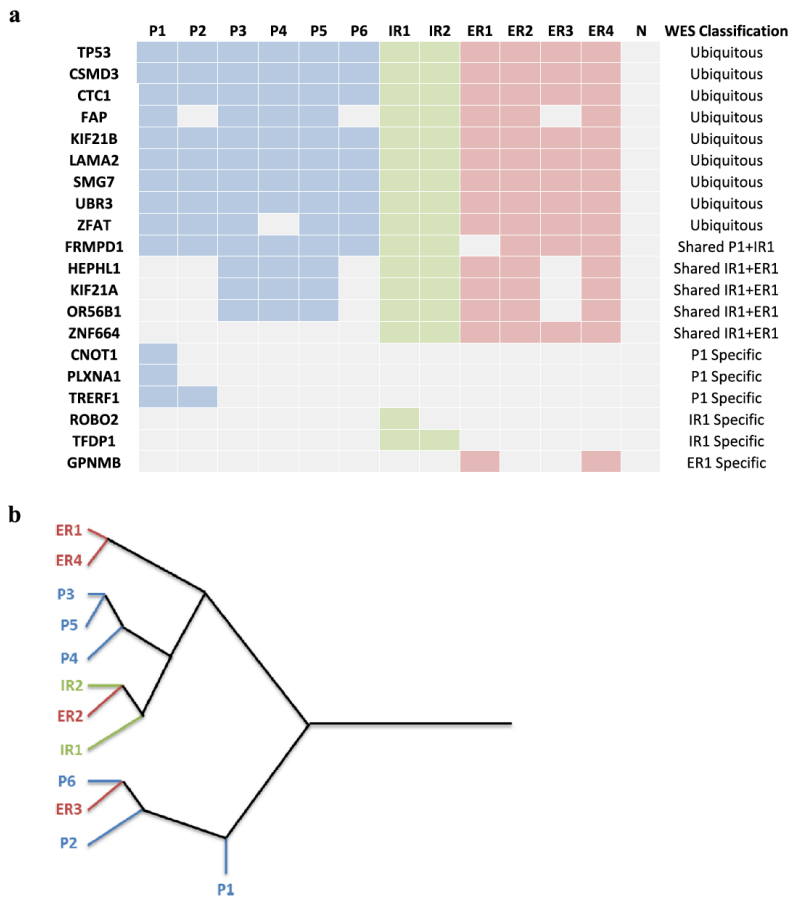


Fig. 4 Intra-tumoral mutational pattern and hierarchical clustering to establish a clonal evolution in a TP53 null HGSC. **a** Selected variants detected by whole-exome sequencing were reanalysed in 12 tumoral samples (P: Primary tumor, IR: intrapelvic recurrence, ER: extrapelvic recurrence) and normal tissue (N) by Sanger sequencing. **b** Hierarchical clustering of the analysed samples on the basis of the similarities in mutation pattern

Sanger sequencing for some of the mutations detected in the exome analysis. The conclusions of this analysis are consistent with a situation in which the primary tumor is composed by mutationally heterogeneous clones, some of which give rise to the recurrences. This heterogeneity is intrinsic to the primary tumor and therefore not expected to be a consequence of the therapy, consistently with what has been previously reported for platinum-based treatment [14]. The hierarchical clustering of the different tumor regions allows to dissect the tumor clonal evolution. The primary tumor regions P3-P5-P4 are the most closely related to the recurrences, with an ancestor tumor subpopulation that would give rise to these primary regions, both intra-pelvic recurrent locations and most of the extra-pelvic (ER1, ER2 and ER4). However, it is worth noting that while the intra-pelvic recurrence seems to have evolved entirely from this ancestor, there is a region of the extra-pelvic recurrence which is more closely related to the other evolutionary branch of the tumor, which gives rise to P1, P2, P6 and, apparently deriving from the latter, ER3. This fact could be regarded as a reflect of a possibly heterogeneous origin for the extra-pelvic recurrence or either as an artifact of the hierarchical clustering process, which does not take into account the CNVs observed in this tumor. Even if the latter is the case, the extra-pelvic recurrence is composed of two distinct subpopulations, one derived from the ancestor clone that gives rise to the upper evolutionary branch (ER1 and ER4) and the other one derived from the intra-pelvic recurrence (ER2), having acquired the ability to migrate to the extra-pelvic area.

As determined by functional annotation and network analysis, cell adhesion and cell cycle control are clearly

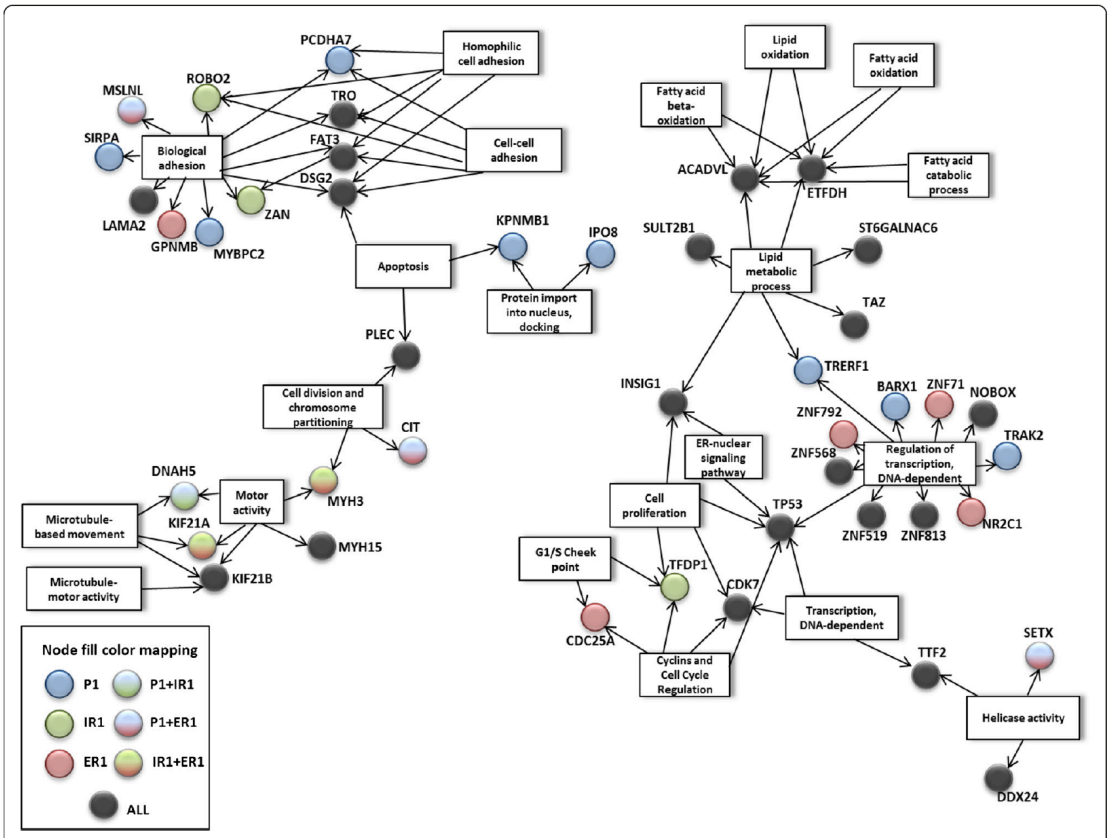


Fig. 5 Network analysis of signaling pathways affected by multiple gene mutation. Functional annotation was performed for variants with a possible negative consequence detected by whole-exome sequencing using David protocols (see M&M). Network analysis by Cytoscape tool revealed signaling pathways containing multiple mutated genes. Functional nodes annotated by Biocarta, Gene Ontology, KEGG or Reactome terms are represented in squared nodes. Mutated genes are represented in circles, which colour indicates the samples in which they are mutated (see node fill color mapping)

affected in this tumor. These functions are intrinsic to tumor growth and spread, and therefore its malfunction could be expected. Some of the genes related to cell adhesion altered in this tumor have been previously reported as related to tumor progression. Down regulation and mutations in *ROBO2* have been reported in prostate, gastric and colorectal cancers [33, 34]; while *LAMA2* down regulation has been reported in drug-resistant ovarian cancer cell lines [35]. Moreover the silencing of *TRO*, coding for trophinin, has been related to cisplatin resistance and increased invasiveness of ovarian cancer cells [36]. Conversely *GPNMB*, a migration-related gene, is frequently overexpressed in triple negative breast cancer among others [37]. Interestingly, *GPNMB* is only mutated in ER1 and ER4, suggesting an important role for this gene in the appearance of extra-pelvic recurrence in this patient. Cell proliferation and cell cycle control are also significantly altered in this tumor, not only by gene mutation but also by copy number changes. *TFDP1*, mutated only in the intra-pelvic recurrences regions, participates in cell cycle control modulating E2F pathway, and mutations in this gene can be found in various cancer databases [38]. Moreover, a recurrent frameshift mutation in *TFDP1* has been reported in colorectal cancer [39]. In addition, 8q11.1q24 region, which is amplified in the three samples analysed by CGH, includes among other genes *MYC*, a key inductor of proliferation which has been shown to be commonly amplified in ovarian cancer [11]. Another important pathway consistently affected in all samples analysed was microtubule-based movement. *KIF21A*, one of these mutated genes, has been shown to be down regulated in a murine lung cancer model due to the aberrant methylation of its promoter [40]. It has also been reported, together with other cytoskeleton-associated proteins, to be implicated in breast, cervix and osteosarcoma cells

survival [41]. Microtubule-based movement is essential for chromosomal partitioning and segregation, an important role in ovarian HGSC, characterised by wide genomic and chromosomal aberrations. This specific case shows extensive copy number variations, with depletion of whole chromosomes 4, 12 and 16. More unexpected was the alteration of lipid metabolism. Mutations related to lipid metabolism were mainly shared by the three samples analysed by WES, suggesting a driver role for this pathway. Some of the mutated lipid-related genes have been reported to be involved in tumor progression in other types of tumors. *ACADVL* expression is down regulated in adrenocortical tumors [42]. Germline polymorphisms in the estradiol-metabolism related enzyme *SULT2B1* have been associated to prostate cancer risk [43], and its expression has been shown to be reduced in this type of tumor [44]. *PCYT2* activity has been shown to be increased in breast cancer cells, enabling them to adapt to metabolic stress [45], and *TRERF1* acts as a cell cycle inhibitor in breast cancer cells through the modulation of progesterone receptor [46]. Altogether, this data supports the hypothesis of a key role for lipid metabolism deregulation in this case of ovarian HGSC. Interestingly, TCGA expression data regarding ovarian cancer shows that the expression of lipid metabolism-related genes is consistently altered in this type of tumor [47]. As a steroid hormone producing organ, steroid metabolism must be tightly regulated in the ovary. Besides, it is well known that lipogenesis is increased in cancerous cells in comparison to normal cells. In addition, lipid catabolism inhibits glycolisis, the major energy source used by malignant cells. Moreover, *TP53* has been shown to be a key regulator of lipid metabolism, inhibiting its anabolism and promoting its catabolism through the regulation of gene expression [48]. Thus, lipid metabolism could represent an important metabolic pathway related to ovarian cancer progression, constituting a new therapeutic target for this type of cancer, frequently resistant to conventional therapy.

Nevertheless it is important to take into consideration that this study has been carried out with a single patient. In order to strengthen its conclusions about the consequences of intra-tumoural heterogeneity in HGSCs, and also to gain a deeper understanding of the biology of p53-null HGSCs, an in-depth study considering a larger number of patients should be undertaken.

Conclusions

Overall, these data shed light on the clonal evolution of the distinct tumor regions and thus on at least some of the genes that may be implicated in its progression and recurrence, and highlights the importance of considering intra-tumor heterogeneity when carrying out genetic and genomic studies, especially when these are aimed to diagnostic procedures or to uncover possible therapeutic strategies.

Additional files

Additional file 1: Supplementary Materials: Whole-exome sequencing: bioinformatics analysis. (PDF 39 kb)

Additional file 2: Figure S1. Fluorescence in situ hybridisation of *PML* gene in primary tumor and recurrence samples shows genomic intra-tumoral heterogeneity. Representative FISH images of *PML* (red) and *RARA* (green, used as control) genes in primary tumor (P1-P6) and recurrence (IR1-IR2 and ER1-ER4) samples. Magnification = 40X. (PDF 115 kb)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AM, JCT, PGS and GMB conceived experiments. AM, ARS, JCT and AMR carried out experiments and analysed data. AM, JCT and ARS created figures. LC, A G-M, JFG, PGS and GMB conceived the study, participated in its design and coordination and wrote the manuscript. All authors were involved in writing the paper and had final approval of the submitted.

Acknowledgements

We thank all the people of Translational Research Laboratory and Immunohistochemical Laboratory from MD Anderson Madrid for their invaluable help. This work has been supported by grants from the AECC network-2012, Telemarató 2013, Instituto de Salud Carlos III (ISCIII) (PI13/00132 and RETIC-RD12/0036/0007), GEIS award 2013, and by the Community of Madrid (S2010/BMD-2303). AM is a predoctoral student supported by FPU fellowship (Spanish Education Ministry). PGS is founded by postdoc contracts from the AECC Scientific Foundation.

Author details

[1]Departamento de Bioquímica, Universidad Autónoma de Madrid (UAM), Instituto de Investigaciones Biomédicas "Alberto Sols" (CSIC-UAM), IdiPAZ, Madrid, Spain. [2]MD Anderson International Foundation, Madrid, Spain. [3]Sistemas Genómicos, Valencia, Spain. [4]Department of Pathology, MD Anderson Cancer Center, Madrid, Spain. [5]Department of Molecular Cytogenetics, MD Anderson Cancer Center, Madrid, Spain. [6]Department of Gynecologic Oncology, MD Anderson Cancer Center, Madrid, Spain. [7]Department of Medical Oncology, MD Anderson Cancer Center, Madrid, Spain.

Received: 28 January 2015 Accepted: 23 November 2015
Published online: 30 November 2015

References

1. Vergote I, Amant F, Kristensen G, Ehlen T, Reed NS, Casado A. Primary surgery or neoadjuvant chemotherapy followed by interval debulking surgery in advanced ovarian cancer. Eur J Cancer. 2011;47:S88–92.
2. Gurung A, Hung T, Morin J, Gilks CB. Molecular abnormalities in ovarian carcinoma: clinical, morphological and therapeutic correlates. Histopathology. 2013;62(1):59–70.
3. Ahmed AA, Etemadmoghadam D, Temple J, Lynch AG, Riad M, Sharma R, et al. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. J Pathol. 2010;221(1):49–56.
4. Yemelyanova A, Vang R, Kshirsagar M, Lu D, Marks MA, Shih Ie M, et al. Immunohistochemical staining patterns of p53 can serve as a surrogate marker for TP53 mutations in ovarian carcinoma: an immunohistochemical and nucleotide sequencing analysis. Mod Pathol. 2011;24(9):1248–53.
5. Hall J, Paul J, Brown R. Critical evaluation of p53 as a prognostic marker in ovarian cancer. Expert Rev Mol Med. 2004;6(12):1–20.
6. Shahin MS, Hughes JH, Sood AK, Buller RE. The prognostic significance of p53 tumor suppressor gene alterations in ovarian carcinoma. Cancer. 2000;89(9):2006–17.
7. Kobel M, Reuss A, Bois A, Kommoss S, Kommoss F, Gao D, et al. The biological and clinical value of p53 expression in pelvic high-grade serous carcinomas. J Pathol. 2010;222(2):191–8.
8. Wojnarowicz PM, Oros KK, Quinn MC, Arcand SL, Gambaro K, Madore J, et al. The genomic landscape of TP53 and p53 annotated high grade ovarian serous carcinomas from a defined founder population associated with patient outcome. PLoS One. 2012;7(9):e45484.

9. Nadkarni NJ, Geest KD, Neff T, Young BD, Bender DP, Ahmed A, et al. Microvessel density and p53 mutations in advanced-stage epithelial ovarian cancer. Cancer Lett. 2013;331(1):99–104.
10. Petitjean A, Achatz MI, Borresen-Dale AL, Hainaut P, Olivier M. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. Oncogene. 2007;26(15):2157–65.
11. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. Nature. 2011;474(7353):609–15.
12. Castellarin M, Milne K, Zeng T, Tse K, Mayo M, Zhao Y, et al. Clonal evolution of high-grade serous ovarian carcinoma from primary to recurrent disease. J Pathol. 2013;229(4):515–24.
13. Zhang J, Shi Y, Lalonde E, Li L, Cavallone L, Ferenczy A, et al. Exome profiling of primary, metastatic and recurrent ovarian carcinomas in a BRCA1-positive patient. BMC Cancer. 2013;13:146.
14. Bashashati A, Ha G, Tone A, Ding J, Prentice LM, Roth A, et al. Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. J Pathol. 2013;231(1):21–34.
15. Khalique L, Ayhan A, Whittaker JC, Singh N, Jacobs IJ, Gayther SA, et al. The clonal evolution of metastases from primary serous epithelial ovarian cancers. Int J Cancer. 2009;124(7):1579–86.
16. Cooke SL, Ng CK, Melnyk N, Garcia MJ, Hardcastle T, Temple J, et al. Genomic analysis of genetic heterogeneity and evolution in high-grade serous ovarian carcinoma. Oncogene. 2010;29(35):4905–13.
17. Khalique L, Ayhan A, Weale ME, Jacobs IJ, Ramus SJ, Gayther SA. Genetic intra-tumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours. J Pathol. 2007;211(3):286–95.
18. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013;31(3):213–9.
19. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012;22(3):568–76.
20. Gonzalez-Perez A, Lopez-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. Am J Hum Genet. 2011;88(4):440–9.
21. Ng PC, Henikoff S. Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet. 2006;7:61–80.
22. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248–9.
23. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science. 1992;258(5083):818–21.
24. Martinez-Ramirez A, Urioste M, Melchor L, Blesa D, Valle L, de Andres SA, et al. Analysis of myelodysplastic syndromes with complex karyotypes by high-resolution comparative genomic hybridization and subtelomeric CGH array. Genes Chromosomes Cancer. 2005;42(3):287–98.
25. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.
26. da Huang W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1–13.
27. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics. 2011;27(3):431–2.
28. Gonzalez-Martin A, Gladieff L, Tholander B, Stroyakovsky D, Gore M, Scambia G, et al. Efficacy and safety results from OCTAVIA, a single-arm phase II study evaluating front-line bevacizumab, carboplatin and weekly paclitaxel for ovarian cancer. Eur J Cancer. 2013;49(18):3831–8.
29. Yuce O, West SC. Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. Mol Cell Biol. 2013;33(2):406–17.
30. Wiedemeyer WR, Beach JA, Karlan BY. Reversing platinum resistance in high-grade serous ovarian carcinoma: targeting BRCA and the homologous recombination system. Front Oncol. 2014;4:34.
31. Wong KK, Izaguirre DI, Kwan SY, King ER, Deavers MT, Sood AK, et al. Poor survival with wild-type TP53 ovarian cancer? Gynecol Oncol. 2013;130(3):565–9.
32. Hentze MW, Kulozik AE. A perfect message: RNA surveillance and nonsense-mediated decay. Cell. 1999;96(3):307–10.
33. Je EM, Gwak M, Oh H, Choi MR, Choi YJ, Lee SH, et al. Frameshift mutations of axon guidance genes ROBO1 and ROBO2 in gastric and colorectal cancers with microsatellite instability. Pathology. 2013;45(7):645–50.
34. Choi YJ, Yoo NJ, Lee SH. Down-regulation of ROBO2 expression in prostate cancers. Pathol Oncol Res. 2014;20(3):517–9.
35. Januchowski R, Zawierucha P, Rucinski M, Zabel M. Microarray-based detection and expression analysis of extracellular matrix proteins in drugresistant ovarian cancer cell lines. Oncol Rep. 2014;32(5):1981–90.
36. Baba T, Mori S, Matsumura N, Kariya M, Murphy SK, Kondoh E, et al. Trophinin is a potent prognostic marker of ovarian cancer involved in platinum sensitivity. Biochem Biophys Res Commun. 2007;360(2):363–9.
37. Rose AA, Grosset AA, Dong Z, Russo C, Macdonald PA, Bertos NR, et al. Glycoprotein nonmetastatic B is an independent prognostic indicator of recurrence and a novel therapeutic target in breast cancer. Clin Cancer Res. 2010;16(7):2147–56.
38. Munro S, Oppermann U, La Thangue NB. Pleiotropic effect of somatic mutations in the E2F subunit DP-1 gene in human cancer. Oncogene. 2014;33(27):3594–603.
39. Chen C, Liu J, Zhou F, Sun J, Li L, Jin C, et al. Next-generation sequencing of colorectal cancers in Chinese: identification of a recurrent frame-shift and gain-of-function indel mutation in the TFDP1 gene. OMICS. 2014;18(10):625–35.
40. Sun W, Iijima T, Kano J, Kobayashi H, Li D, Morishita Y, et al. Frequent aberrant methylation of the promoter region of sterile alpha motif domain 14 in pulmonary adenocarcinoma. Cancer Sci. 2008;99(11):2177–84.
41. Groth-Pedersen L, Aits S, Corcelle-Termeau E, Petersen NH, Nylandsted J, Jaattela M. Identification of cytoskeleton-associated proteins essential for lysosomal stability and survival of human cancer cells. PLoS One. 2012;7(10):e45381.
42. Soon PS, Libe R, Benn DE, Gill A, Shaw J, Sywak MS, et al. Loss of heterozygosity of 17p13, with possible involvement of ACADVL and ALOX15B, in the pathogenesis of adrenocortical tumors. Ann Surg. 2008;247(1):157–64.
43. Levesque E, Laverdiere I, Audet-Walsh E, Caron P, Rouleau M, Fradet Y, et al. Steroidogenic germline polymorphism predictors of prostate cancer progression in the estradiol pathway. Clin Cancer Res. 2014;20(11):2971–83.
44. Seo YK, Mirkheshti N, Song CS, Kim S, Dodds S, Ahn SC, et al. SULT2B1b sulfotransferase: induction by vitamin D receptor and reduced expression in prostate cancer. Mol Endocrinol. 2013;27(6):925–39.
45. Zhu L, Bakovic M. Breast cancer cells adapt to metabolic stress by increasing ethanolamine phospholipid synthesis and CTP:ethanolaminephosphate cytidylyltransferase-Pcyt2 activity. Biochem Cell Biol. 2012;90(2):188–99.
46. Gizard F, Robillard R, Gross B, Barbier O, Revillion F, Peyrat JP, et al. TReP-132 is a novel progesterone receptor coactivator required for the inhibition of breast cancer cell growth and enhancement of differentiation by progesterone. Mol Cell Biol. 2006;26(20):7632–44.
47. Ying H, Lv J, Ying T, Jin S, Shao J, Wang L, et al. Gene-gene interaction network analysis of ovarian cancer using TCGA data. J Ovarian Res. 2013;6(1):88.
48. Goldstein I, Rotter V. Regulation of lipid metabolism by p53 - fighting two villains with one sword. Trends Endocrinol Metab. 2012;23(11):567–75.

Modern Pathology (2016), 1–12

Genetic analysis of uterine aspirates improves the diagnostic value and captures the intra-tumor heterogeneity of endometrial cancers

Alba Mota[1,2], Eva Colás[3], Pablo García-Sanz[1,2], Irene Campoy[4], Alejandro Rojo-Sebastián[5], Sonia Gatius[3], Ángel García[6], Luis Chiva[5], Sonsoles Alonso[5], Antonio Gil-Moreno[4,7], Xavier González-Tallada[3], Berta Díaz-Feijoo[4,7], August Vidal[8], Patrycja Ziober-Malinowska[9], Marcin Bobiński[9], Rafael López-López[10], Miguel Abal[10], Jaume Reventós[8,11], Xavier Matias-Guiu[3,8] and Gema Moreno-Bueno[1,2]

[1]Departamento de Bioquímica, Universidad Autónoma de Madrid (UAM), Instituto de Investigaciones Biomédicas 'Alberto Sols' (CSIC-UAM), IdiPAZ, Madrid, Spain; [2]MD Anderson International Foundation, Madrid, Spain; [3]Department of Pathology and Molecular Genetics and Research Laboratory, Hospital Universitari Arnau de Vilanova, University of Lleida, IRBLLEIDA, Lleida, Spain; [4]Biomedical Research Group in Gynaecology, Vall d'Hebron Research Institute and Hospital, and Universitat Autonoma de Barcelona, Barcelona, Spain; [5]MD Anderson Cancer Center Madrid, Madrid, Spain; [6]Department of Pathology, Hospital Universitario Vall Hebron, Barcelona, Spain; [7]Department of Obstetrics and Gynaecology, Hospital Universitari Vall d'Hebron, Barcelona, Spain; [8]Hospital Universati de Bellvitge, Institut d'Investigacions Biomèdiques de Bellvitge (IDIBELL), Barcelona, Spain; [9]First Chair and Department of Gynaecological Oncology and Gynaecology, Medical University of Lublin, Lublin, Poland; [10]Translational Medical Oncology, Health Research Institute of Santiago (IDIS), SERGAS, Santiago, Spain and [11]Departament de Ciències Bàsiques, Universitat Internacional de Catalunya, Barcelona, Spain

Endometrial cancer is the most common cancer of the female genital tract in developed countries. Although the majority of endometrial cancers are diagnosed at early stages and the 5-year overall survival is around 80%, early detection of these tumors is crucial to improve the survival of patients given that the advanced tumors are associated with a poor outcome. Furthermore, correct assessment of the pre-clinical diagnosis is decisive to guide the surgical treatment and management of the patient. In this sense, the potential of targeted genetic sequencing of uterine aspirates has been assessed as a pre-operative tool to obtain reliable information regarding the mutational profile of a given tumor, even in samples that are not histologically classifiable. A total of 83 paired samples were sequenced (uterine aspirates and hysterectomy specimens), including 62 endometrioid and non-endometrioid tumors, 10 cases of atypical hyperplasia and 11 non-cancerous endometrial disorders. Even though diagnosing endometrial cancer based exclusively on genetic alterations is currently unfeasible, mutations were mainly found in uterine aspirates from malignant disorders, suggesting its potential in the near future for supporting the standard histologic diagnosis. Moreover, this approach provides the first evidence of the high intra-tumor genetic heterogeneity associated with endometrial cancer, evident when multiple regions of tumors are analyzed from an individual hysterectomy. Notably, the genetic analysis of uterine aspirates captures this heterogeneity, solving the potential problem of incomplete genetic characterization when a single tumor biopsy is analyzed.
Modern Pathology advance online publication, 2 September 2016; doi:10.1038/modpathol.2016.143

Correspondence: Dr G Moreno-Bueno, PhD, Departamento de Bioquímica, Facultad de Medicina (UAM), Instituto de Investigaciones Biomedicas 'Alberto Sols' CSIC-UAM, C/Arzobispo Morcillo 2, Madrid 28029, Spain or MD Anderson International Foundation, C/ Gomez Hemans 2, Madrid 28033, Spain.
E-mail: gmoreno@iib.uam.es or gmoreno.fundacion@mdanderson.es
Received 20 December 2015; revised 29 June 2016; accepted 4 July 2016; published online 2 September 2016

www.modernpathology.org

Endometrial cancer is the fourth most common cancer among women in developed countries and the most frequent cancer of the female genital tract.[1] Endometrial cancer is mainly classified into two groups with different clinical, pathological, and molecular features.[2,3] Type I or endometrioid endometrial carcinomas are normally low-grade,

estrogen-related tumors with a good prognosis. These tumors are the most common endometrial cancers and they usually arise in perimenopausal women, preceded by or coexisting with endometrial hyperplasia. Type II tumors are high-grade non-endometrioid endometrial carcinomas, unrelated to estrogen, which occur in older women and have a poor prognosis.[4,5] The majority of non-endometrioid endometrial carcinomas are serous endometrial carcinomas, although there are also less frequent histological subtypes, such as uterine carcinosarcomas (also known as malignant mixed müllerian tumors), uncommon biphasic neoplasms with malignant epithelial elements, and a sarcomatoid component.[5] At the molecular level, significant differences are evident between type I and II carcinomas. Mutations in *PTEN*, *PIK3CA*, *PIK3R1*, *KRAS*, *CTNNB1*, *FGFR2*, and *ARID1A* are associated with endometrioid endometrial carcinomas, whereas mutations in *TP53*, *PIK3CA*, and *PPP2R1A* are frequent in non-endometrioid endometrial carcinomas.[6,7] In addition, microsatellite instability is evident in one-third of endometrioid endometrial carcinomas, a feature that is unusual in non-endometrioid endometrial carcinomas, which are more commonly characterized by chromosomal instability.[8,9]

The majority of endometrial cancers are diagnosed at early stages and the associated 5-year overall survival is around 80%. Nevertheless, the survival rate decreases to 57–46% for high-grade tumors.[5] Furthermore, carcinosarcomas account for a high percentage of mortality despite constituting only 5–6% of endometrial cancers, principally because 60% of the patients presents extra uterine disease at the moment of diagnosis. In these cases more than 50% will suffer recurrence after surgery and adjuvant treatment.[10,11]

As with many other tumors, early detection of endometrial cancer is crucial to increase patient survival, particularly as advanced tumors are associated with a worse outcome. Moreover, the correct assessment of pre-clinical diagnosis is also decisive, as this will guide the pre-operative and surgical management of the patient.[12] In this sense, the use of uterine aspirates (Pipelle biopsies) as diagnostic pre-operative biopsies is widely recommended, representing a minimally invasive and highly sensitive procedure. However, the failure rate in obtaining such samples is around 8%, whereas 13% of the samples turn out to be histologically inadequate, figures that are significantly higher in postmenopausal women.[12,13] Moreover, discrepancies between pre- and post-operative biopsies have been observed with respect to histological grade, which could lead to a misclassification and the use of inappropriate therapeutic strategies.[14,15] As such, there has been some interest in identifying molecular markers in uterine aspirates, enhancing their potential as a diagnostic sample for both histological classification and molecular characterization of tumors.[16–18]

Describing the genetic profile of tumors can be decisive for their accurate diagnosis and for therapeutic decision-making. However, intra-tumor genetic heterogeneity represents a challenge that hampers the correct characterization of tumor samples.[19] The current study reveals how uterine aspirates are a potentially useful tool to circumvent the problems derived from intra-tumor heterogeneity when genetically characterizing endometrial cancer. We defined the mutational profile of endometrial cancers in paired pre-operative uterine aspirates and hysterectomy specimens from patients. The data obtained not only confirmed the utility of these aspirates to detect the mutations in primary tumors, even when a pathological diagnosis could not be achieved by other means, but importantly, they also reflected the high intra-tumor genetic heterogeneity found in endometrial cancers. These results show that the genetic analysis of uterine aspirates provides information that the pathologist may find useful to reduce the rate of false-negative diagnoses. In summary, we show the importance of uterine aspirates in studying endometrial cancer at the molecular level, supporting the potential of non-invasive biopsies for the diagnosis and characterization of certain tumor types.

Materials and methods

Sample Description

A total of 62 endometrial cancer cases (44 endometrioid endometrial carcinomas, 9 serous endometrial carcinomas, 9 carcinosarcomas) were collected at Vall d'Hebron Hospital (Barcelona), Arnau de Vilanova University Hospital (Lleida), MD Anderson Cancer Center (Madrid) and Medical University (Lubin) between 2010 and 2015. The median age of the patients was 67 (±12, endometrioid endometrial carcinomas), 75 (±8, serous endometrial carcinomas), and 72 (±8, carcinosarcomas) and the histopathological data of the tumors studied can be found in Supplementary Table 1. Endometrial tissue from endometrial aspirates and hysterectomy specimens were analyzed from each subject. In addition, samples obtained from 10 patients diagnosed with atypical hyperplasia, were collected at Hospital Universitari de Bellvitge and used as an example of precursor malignant neoplasia. A total of 27 patients not diagnosed with cancer were also analyzed as controls for the studies of the mutational profile (7 non-atypical hyperplasia endometrium, the endometrium from 7 patients with leiomyoma, and 13 normal endometrium). In 11 of these controls uterine aspirates and their respective hysterectomy specimen (endometrial tissue) were analyzed, whereas in the remainder only a uterine aspirate was available. Uterine aspirates were collected using a Pipelle de Cornier to obtain the sample that was then centrifuged for 20 min, as described previously.[16] The pellet containing the cells from the uterine cavity was processed as formalin-fixed and paraffin-embedded tissue for further DNA extraction.

A second uterine aspirate was obtained in the operating room just before surgery, being the tumor material frozen at −80 °C for hematoxylin and eosin stain examination and DNA extraction. Only in which the formalin-fixed and paraffin-embedded uterine aspirate material was used up in the histologic analysis, frozen tissue was used for the study. The study was approved by the local ethical committee from each institution, and a complete written informed consent was obtained from all patients.

DNA Extraction and Mutational Analysis

DNA was obtained from formalin-fixed paraffin-embedded and frozen samples using phenol extraction and ethanol precipitation, and 10 ng were used for sequencing. Multiplex PCR to prepare amplicon libraries was performed using the Ion AmpliSeq Library Kit 2.0 and Ion AmpliSeq Cancer Hotspot Panel v2 (Life Technologies). For PCR, a total of 17 and 20 cycles were used for the frozen and formalin-fixed paraffin-embedded samples, respectively. The PCR template preparation and enrichment were performed using Ion PGM Template OT2 200 Kit and the Ion OneTouch 2 System. Finally, the Ion PGM Sequencing 200 Kit v2 and Ion PGM System (Life Technologies) were used for DNA sequencing according to the manufacturer's protocols. Duplicates were analyzed for 10 % of the samples, rendering equivalent results. For the bioinformatics analysis, see Supplementary Methods.

Sanger Sequencing

To validate the mutations, a total of 88 of the 476 variants found in the samples analyzed were Sanger sequenced. The PCR conditions and amplicon lengths used are indicated in Supplementary Table 2. Only 6 of these variants were not confirmed by Sanger sequencing, which was probably due to their poor quality and/or their frequencies below 10% in the Ion PGM sequencing analysis (Supplementary Table 3).

Statistical Analysis

A paired *t*-test was used to compare the data from the hysterectomy tumor samples and uterine aspirates. Two-tailed tests were performed and 95% confidence intervals (CIs) were accepted. The mutation discovery rate was calculated in each sample (aspirate or tumor region) from the same patient according to the following equation:

$$\frac{\sum \text{Sample mutation}}{\sum(\sum \text{Aspirate mutation} + \sum \text{Tumor region 1 mutation} + \ldots + \sum \text{Tumor region } n \text{ mutation})} \times 100$$

The Pearson coefficient was used to analyze the correlation between the percentage of tumor cells in patient samples and the MDR. P values < 0.05 were considered statistically significant and the statistical analyses were performed using the SPSS Statistics 17.0 software (SPSS, Chicago, IL, USA).

Results

Identification of the Mutational Profile in Paired Uterine Aspirate and Hysterectomy Specimen Samples

Uterine aspirates are thought to be highly sensitive and specific biopsies for the pre-operative diagnosis of endometrial cancer, especially when based on biomarker expression.[13,16–18] To investigate the usefulness of mutation detection in uterine aspirates, the molecular profile of paired samples (pre-operative uterine aspirates and the corresponding resected surgical specimen) from 54 patients with endometrial cancer (37 endometrioid endometrial carcinomas, 9 serous endometrial carcinomas, and 8 carcinosarcomas) and 10 patients with atypical hyperplasia was analyzed using AmpliSeq Cancer Hotspot Panel v2. This panel analyzes approximately 2800 cancer mutations of 50 oncogenes and tumor suppressor genes, some of which are frequently altered in endometrial cancer (*PTEN*, *KRAS*, *FGFR2*, *CTNNB1*, *PIK3CA*, *FBXW7*, and *TP53*). In addition, a total of 27 patients not diagnosed with cancer were also analyzed as control cases (7 cases of non-atypical hyperplasia, 7 cases with leiomyomas and 13 with a normal endometrium).

Sequencing analysis revealed the presence of mutations in 51 of the 54 aspirates from cancer patients (Table 1; Supplementary Table 3A) and in 5 of the 10 aspirates from atypical hyperplasia cases (Supplementary Table 4). By contrast, mutations were only identified in 1 of the 27 control patients (data not shown). Although it is currently unfeasible to reach a diagnosis of endometrial malignancies based exclusively on genetic alterations, these results indicate that genetic analysis of uterine aspirates may offer reliable support to histological diagnosis.

Mutations identified in the different subgroups of patients were consistent with previous studies.[6,7] In summary, endometrioid endometrial carcinomas carried mutations in *PTEN* (71.1% of patients), *PIK3CA* (39.5%), *CTNNB1* (28.9%), *TP53* (28.9%), *FGFR2* (23.7%), *KRAS* (21.1%), and *CDKN2A* (10.5%). In addition, we also detected mutations in less commonly affected genes, such as: *ABL1*, *AKT1*, *APC*, *ATM*, *BRAF*, *ERBB2*, *FBXW7*, *KIT*, *RB1*, and *VHL1* (5.3%); and *GNA11*, *GNAS*, *HNF1A*, *MET*, *MLH1*, *NRAS*, *RET*, *STK11*, *SMAD4*, *SMARCB1*, and *SMO* (2.6%). As expected, the most frequently mutated gene in serous carcinomas and carcinosarcomas samples was *TP53* (77.7 and 87.5%, respectively). Frequencies found in our series were generally higher than those detected in the The Cancer Genome Atlas dataset[6] (Supplementary Figure 1A). This could be explained taking into

Table 1 Summary of the histological grade and mutational profile in endometrial cancers and their paired uterine aspirates

Patient	*Aspirate grade*	*Hysterectomy grade*	*Common variants (aspirate and hysterectomy)*[a]	*Hysterectomy variants (not detected in aspirate)*[b]	*Aspirate variants (not detected in hysterectomy)*	*% Hysterectomy variants detected in aspirate*[(a/(a+b))]
EEC-1	2	3	*PTEN* (2), *TP53* (2), *APC*			100
EEC-2	3	3	*PTEN, CTNNB1, CDKN2A*			100
EEC-3	3	3	*PTEN, PIK3CA* (2), *ABL1*			100
EEC-4	1	1	*PTEN, KRAS*		*PIK3CA* (3)	100
EEC-5	1	1	*PTEN, APC*			100
EEC-6	1	3	*FGFR2, PIK3CA, KIT*		*PIK3CA*	100
EEC-7	1	3	*FGFR2, PTEN* (2), *TP53*	*PTEN, TP53, SMARCB1, CTNNB1, CKN2A*		44.4
EEC-8	1	3	*PIK3CA, CTNNB1*			100
EEC-9	2	3	*FGFR2, PTEN, PIK3CA, CTNNB1*			100
EEC-10	3	3	*PIK3CA*			100
EEC-11	1	1	*FGFR2, FBXW7*			100
EEC-12	2	2	*PTEN* (3), *KRAS, RB1, ERBB2, TP53, PIK3CA, CTNNB1, FBXW7*			100
EEC-13	1	3	*NRAS, PTEN* (2), *ATM, HNF1A, PIK3CA, SMO, ABL1, CDKN2A*	*KRAS, GNA11* (2)	*PTEN, ATM, TP53* (2), *SMAD4, GNAS, CTNNB1*	75
EEC-14	2	3	*PTEN* (2), *ERBB2*			100
EEC-15	1	1	*PTEN*	*RET, STK11, PIK3CA*	*VHL*	25
EEC-16	2	3	*PTEN*(3), *TP53*			100
EEC-17	2	3	*KRAS, AKT1*			100
EEC-18	1	1	*PTEN, PIK3CA*			100
EEC-19	2	3		*FGFR2, PTEN* (3), *KRAS*		0
EEC-20	2	3	*KRAS*			100
EEC-21	3	1			*TP53, PIK3CA*	—
EEC-22	2	3	*PTEN* (2)			100
EEC-23	1	3	*FGFR2*			100
EEC-24	1	3	*PTEN, PIK3CA*		*CTNNB1, CDKN2A*	100
EEC-25	2	3	*TP53*			100
EEC-26	2	3	*KRAS, TP53*			100
EEC-27	1	3	*FGFR2, MLH1*	*PTEN*		66.7
EEC-28	2	2	*PTEN* (2), *CTNNB1*			100
EEC-29	3	2	*PTEN*			100
EEC-30	3	3	*PTEN, TP53, PIK3CA*			100
EEC-31	2	2	*PTEN*			100
EEC-32	2	2	*PTEN, CTNNB1*			100
EEC-33	1	1	*FGFR2, AKT1, CTNNB1*			100
EEC-34	2	2	*FGFR2, PTEN*			100
EEC-35	1	3	*PTEN* (3), *KRAS, PIK3CA*	*ATM, RB1, TP53, MET*	*RB1, KIT*	55.6
EEC-36	1	3	*PTEN, BRAF*			100
EEC-37	1	3	*PTEN, PIK3CA* (2)			100
SEC-1	3	3	*TP53*			100
SEC-2	3	3	*TP53*			100
SEC-3	3	3	*KRAS*			100
SEC-4	3	3		*ABL1*	*TP53*	0
SEC-5	3	3	*TP53, PIK3CA, BRAF, ATM*			100
SEC-6	3	3		*TP53*		0
SEC-7	3	3	*PIK3CA*			100
SEC-8	3	3		*TP53 (2), FBXW7*		0
SEC-9	3	3	*TP53*			100
CS-1	3	3	*PTEN, TP53, PIK3CA*	*FBXW7*	*KRAS*	75
CS-2	3	3	*TP53*			100
CS-3	3	3	*TP53*			100
CS-4	3	3	*TP53*			100
CS-5	3	3	*TP53*			100
CS-6	3	3	*PTEN, KRAS, PIK3CA*			100
CS-7	3	3	*KRAS, TP53*	*IDH2, TP53 (3), EGFR*		28.6
CS-8	3	3	*TP53*			100

[a]Number of common mutations (aspirate and hysterectomy).
[b]Number of hysterectomy variants (not detected in aspirate).

account the sequencing method applied in each study. Whereas The Cancer Genome Atlas study[6] performed whole-exome sequencing (mean coverage around to 50×) our study has been developed with targeted sequencing (mean coverage around to 1000×), allowing to detect more accurately the mutations, specially those with low frequency.[20] *TP53* mutation frequency was particularly high in our series, probably due to the presence of mutations in 8 of the 24 high-grade endometrioid carcinomas (Supplementary Figure 1B). To be sure that these cases were not misclassified a second pathology review was performed, confirming the initial diagnosis (Supplementary Table 5).

To gain further insight into the suitability of uterine aspirates to detect mutations and consequently, to estimate the potential of uterine aspirates to characterize endometrial caner from a genetic point of view, we analyzed the percentage of pathogenic variants present in hysterectomy specimens that were also detected in aspirates (Figure 1). All the mutations detected in the surgical tumor tissue were also found in 30 out of the 36 aspirates (83.3%) from endometrioid endometrial carcinoma patients. In terms of the rest of the samples, 50–75% of the mutations detected in the hysterectomy specimen also appeared in the corresponding aspirate in three of them (8.3% of the total), although in two of them (5.6% of the total) the aspirate contained 25–50% of the mutations present in the surgical tissue. Only in 1 patient did we fail to detect any of the mutations identified in the hysterectomy sample in the corresponding aspirate, accounting for 2.8% of the total cases. Conversely, in 1 other patient mutations were detected in the uterine aspirate, whereas none were identified in the surgical sample (Table 1). Furthermore, in 6 of the 9 uterine aspirates from serous carcinomas patients 100% of the mutations identified in the hysterectomy specimens were detected in the aspirate (66.7% of the total cases), as in 6 of the 8 carcinosarcoma cases (75% of the total).

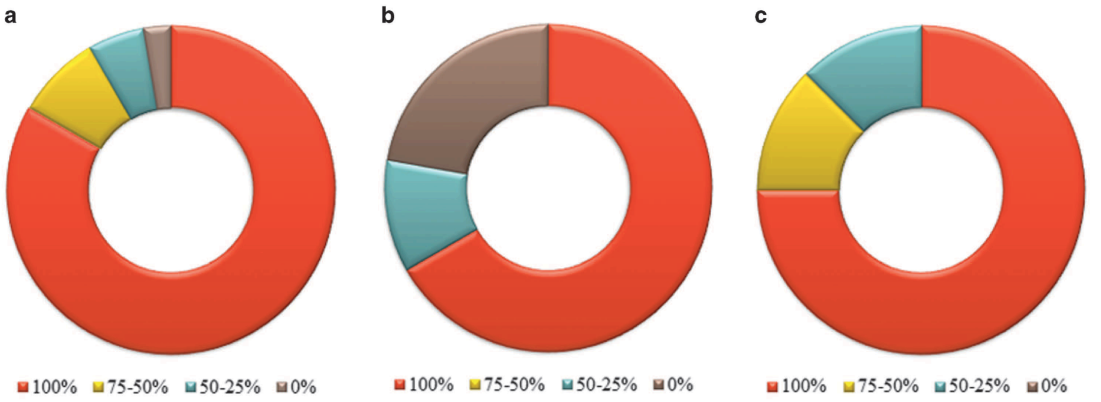
Misclassifying the histological grade of pre-operative biopsies can have grave consequences[14,15,18] and indeed, in our samples 22 of the 37 endometrioid endometrial carcinomas uterine aspirates (59.5%) were misclassified with respect to their grade during the pathological diagnosis, the majority of them being attributed with a lower grade than that detected in the definitive hysterectomy specimen (Table 1). However, in 17 of these 22 (77.2%) discordant classifications, the uterine aspirates were concordant in the mutational analysis, showing all the mutations detected in their respective surgical specimen. Nevertheless, no relationship between mutational status and histological type or grade has been previously described, and nor was one found in our series. Consequently, these results confirm that the genetic analysis of uterine aspirates as a pre-operative biopsy can reliably reproduce the molecular status of the tumor in a pre-clinical setting. However, further studies into the mutational profile and histological grade will be necessary to take the mutational information from uterine aspirates into account when assessing the tumor grade.

Genetic Analysis Helps to Reduce the Rate of False-Negative Diagnoses in Uterine Aspirates

The histologic analyses of uterine aspirates fail to distinguish the presence or absence of malignancy in around 13% of the cases, either due to the small proportion of representative tumor cells or to the poor quality of the specimen.[21–23] To further investigate the potential of genetic analysis of uterine aspirates as an informative tool for endometrial cancer diagnosis, we assessed the tumor mutations that could be detected in uterine aspirates that could not be evaluated on a pathological basis (Figure 2a). Mutational analysis was performed on eight paired samples of non-diagnosable uterine aspirates from patients who turned out to have endometrial cancer and on the corresponding hysterectomy specimens (7 endometrioid endometrial carcinomas and 1 carcinosarcoma). Interestingly, seven of the eight non-evaluable uterine aspirates had a similar mutation profile to that of their paired surgical sample (Figure 2b; Supplementary Table 3B). We did not find mutations in the uterine aspirate from one patient, as was also the case in the paired hysterectomy tumor tissue. We validated these results by Sanger sequencing and immunohistochemistry when material was available. For example, the *CTNNB1* mutation in case EEC-38 was validated by Sanger sequencing in the aspirate and surgical tissue (Figure 2c). To validate this, we also analyzed β-catenin expression by immunohistochemistry in surgical tissue (Figure 2d). These results demonstrated that genetic sequencing complements pathological analysis and contributes significantly to a more comprehensive characterization of the tumor at very early stages of diagnosis, providing valuable information for its correct classification.

Genetic Analysis of Uterine Aspirates Captures the Intra-Tumor Heterogeneity Found in Endometrial Carcinomas

It is well known that human cancers display substantial intra-tumor heterogeneity, not only in cellular morphology or gene expression but also in terms of genetic variation.[24,25] This phenomenon represents an important challenge to accurate diagnosis and therapeutic decision-making.[19] Although recent studies showed intra-tumor genetic heterogeneity in gynecological cancers like ovarian cancer,[26,27] heterogeneity at the mutational level has not been described in endometrial cancer to our knowledge. Interestingly, the comparison between the mutational profile of uterine aspirates and hysterectomy specimens highlighted the presence of additional mutations in 11 out of the 54 uterine



Hysterectomy specimen mutations detected in uterine aspirate (%)	Endometrioid endometrial carcinomas n (%)
100	30 (83)
75-50	3 (8)
50-25	2 (6)
0	1 (3)

Hysterectomy specimen mutations detected in uterine aspirate (%)	Serous endometrial carcinomas n (%)
100	6 (67)
75-50	0 (0)
50-25	1 (11)
0	3 (22)

Hysterectomy specimen mutations detected in uterine aspirate (%)	Carcinosarcomas n (%)
100	6 (75)
75-50	1 (12)
50-25	1 (12)
0	0 (0)

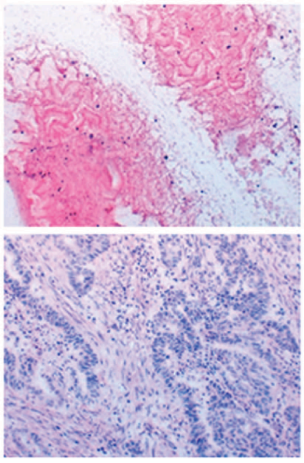
Figure 1 Percentage of mutations in hysterectomy specimens identified in paired uterine aspirate. Graphs represents the percentage (100%, 75–50%, 50–25%, or 0%) of the mutations found in surgical tumor samples and paired aspirates in endometrioid carcinoma (**a**), serous carcinoma (**b**), and carcinosarcoma (**c**) samples.

aspirates, mutations that were not present in the corresponding paired surgical tissue (Table 1). Thus, we examined whether these differences might reflect the intra-tumor genetic heterogeneity in this clinical context.

To explore this hypothesis, genetic sequencing analysis was performed on additional tumor regions from 21 of the endometrial cancer hysterectomy specimens previously studied (14 endometrioid endometrial carcinomas, 5 serous endometrial carcinomas and 5 carcinosarcomas: Table 2). Comparative mutation analysis revealed differences in the mutational profiles of the distinct regions of the endometrioid endometrial carcinomas tumor tissue analyzed from 10 out of 14 patients (71.4%), confirming the presence of intra-tumor heterogeneity (Supplementary Figure 2; Supplementary Table 3C). For example, in the three different tumor regions analyzed from case EEC-1 (Figure 3a), a total of 5 mutations in *PTEN*, *TP53*, and *APC* were detected, with one of the regions (tumor region 1) carrying all five, whereas the other two (tumor region 2 and 3) carried 2 and 3 mutations, respectively. In the remaining cases (4/14, 28.6%), a similar mutational profile was seen in all the samples analyzed (Supplementary Figure 2; Supplementary Table 3C), suggesting that these cases did not harbor significant intra-tumor heterogeneity, at least with respect to the genes and tumor regions studied. For example, this was the case of patient EEC-11 from whom all the samples analyzed carried mutations in *FGFR2* and *FBXW7* (Figure 3b). In contrast to the endometrioid endometrial carcinomas, intra-tumor heterogeneity was only detected in 1 of the 5 (20%) serous carcinomas and 1 of the 5 (20%) carcinosarcomas when additional tumor regions were analyzed (Supplementary Figure 3). The low proportion of mutational heterogeneity in cases with serous and carcinosarcoma histology could be due to the fact that chromosomal instability is a more frequent molecular alteration than punctual genetic changes in these tumor types,[9] a modification that cannot be properly detected with the sequencing platform used here.

The sensitivity of mutation detection in each sample was scored as the mutation discovery rate, which indicates the proportion of mutations detected in each sample with respect to the total mutations observed in all the samples studied from a given patient (see 'Materials and methods' section). The mutation discovery rate was significantly higher in the endometrioid uterine aspirates than in the matched surgical tumor tissue, with a mean of 94.1% for uterine aspirates and 77.2% for individual hysterectomy tissue samples. This difference increased when low-quality mutations were not considered, decreasing the mutation discovery rate for surgical tumor samples to 67.5%, whereas the mutation discovery rate of the aspirates remained unaltered (Figure 3c). However, no significant differences were found in the serous carcinoma or carcinosarcoma samples. Differences in the mutation

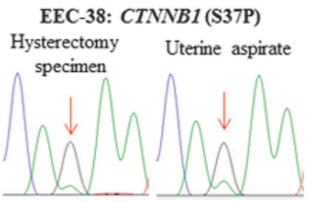
a



b

Patient	Common variants (Hysterectomy and Aspirate)	Hysterectomy variants (not detected in Aspirate)	Aspirate variants (not detected in Hysterectomy)
EEC-38	*CTNNB1*(S37P)	*TP53(P98L), VHL (Q164*)*	-
EEC-39	-	-	-
EEC-40	*FBXW7* (R505H), *PTEN* (Frameshift), *JAK3* (V722I)	-	-
EEC-41	*PTEN* (R130G), *FGFR2* (S252W)	*FGFR2* (Y378C)	-
EEC-42	*PTEN* (R130G)	*PIK3CA* (K111E)	-
EEC-43	*PIK3CA* (E542K), *PTEN* (R130G), *JAK3* (V722I)	-	*TP53* (Frameshift)
EEC-44	*PTEN* (G129R), *CTNNB1* (D32H)	-	*KRAS* (G12V)
CS-9	*FGFR2* (S252W), *PTEN* (Q245*, E284*)	-	-

c



d

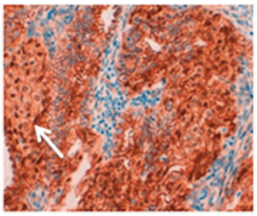


Figure 2 Genetic analysis of non-evaluable uterine aspirates. Paired samples of non-evaluable uterine aspirates and hysterectomy specimens were analyzed genetically. (**a**) Representative hematoxylin and eosin image of a uterine aspirate (upper image) and its paired surgical sample (lower image). (**b**) Summary of the mutations detected in the paired uterine aspirate and hysterectomy samples. Analysis of *CTNNB1* (*β*-catenin) mutation (S37P) found in patient EEC-38 by (**c**) the Sanger sequencing in hysterectomy specimen and uterine aspirate samples and by (**d**) immunohistochemistry analysis. The white arrow label the nuclear localization of *β*-catenin, which is suggestive of mutations (panel magnification × 20).

discovery rate are mainly found in heterogeneous tumors, due to the differences observed in the mutational profile between each tumor region (Figure 4). In 8 of the 10 (80%) heterogeneous endometrioid tumors, uterine aspirates reflected a higher mutation discovery rate than the tumor region used for the pathological diagnosis (tumor region 1). Only in one patient (EEC-7) the mutation discovery rate of the uterine aspirate was lower than that for the diagnostic tumor region, although it was equal or higher than that derived from the two other regions from that patient. In fact, the mutation discovery rate value was higher in uterine aspirates than in at least one tumor region in all cases where there was tumor heterogeneity. These differences seem not to be related to the proportion of the tumor tissue in each region analyzed as there was no significant correlation in a Pearson test (data not shown). These data confirm that genetic analysis of uterine aspirates detects a more representative mutational landscape of the tumor, reproducing in a single sample the intra-tumor heterogeneity found in the different tumor regions.

Discussion

Advances in next-generation sequencing have revealed that genetic heterogeneity must be taken into account to fully understand tumor biology.[19,28,29] Indeed, over and above the inter-patient heterogeneity,[30] intra-tumor heterogeneity represents a real challenge for the precise characterization and adequate management of tumors.[19,31] The presence of different cell populations within a tumor with specific genomic, genetic and/or epigenetic characteristics has been demonstrated in numerous tumor types, including solid tumors and hematologic malignancies.[19] Indeed, intra-tumor heterogeneity has been observed among gynecological cancers, particularly in high-grade serous ovarian carcinomas,[27,32–34] although this issue has not been studied in endometrial cancer so far. Therefore, a

Table 2 Endometrial cancer cases studied in the intra-tumor heterogeneity analysis

Patient	*Tumor regions analyzed*	*Total variants detected in tumor regions*	*Total variants detected in uterine aspirates*	*Intra-tumor heterogeneity*
EEC-1	3	5	5	Yes
EEC-2	2	3	3	Yes
EEC-3	3	4	4	Yes
EEC-4	4	2	5	Yes
EEC-5	3	2	2	Yes
EEC-6	3	3	4	Yes
EEC-7	4	9	4	Yes
EEC-8	4	2	2	No
EEC-9	4	4	4	Yes
EEC-10	4	1	1	No
EEC-11	4	2	2	No
EEC-12	4	10	9	Yes
EEC-13	3	12	16	Yes
EEC-14	4	3	3	No
SEC-1	3	1	1	No
SEC-2	3	1	1	No
SEC-3	3	1	1	No
SEC-4	3	1	1	Yes
SEC-5	3	4	4	No
CS-1	4	4	4	Yes
CS-2	3	1	1	No
CS-3	3	1	1	No
CS-4	3	1	1	No
CS-5	2	1	1	No

better understanding of the genetic heterogeneity underlying the biological and phenotypic evolution of endometrial is crucial to understand the clinical behavior of this disease. In this sense, the majority of the endometrioid carcinomas analyzed here have variable mutational profiles in the different tumor regions. By contrast, only 20% of serous carcinomas and 20% of carcinosarcomas showed mutational heterogeneity, which perhaps reflects the more frequent genetic mutations in endometrioid than in serous and carcinosacomas,[6] the latter more often displaying large genomic changes.[9] Therefore, a genomic study should be carried out on these tumor types to define the implication of copy number variation in intra-tumor heterogeneity, as previously described in high-grade serous ovarian carcinomas.[27,33]

Intra-tumor clonal heterogeneity is thought to influ;ence therapeutic resistance and tumor progression,[35] with some studies suggesting that some clones are genetically predisposed to resist therapy.[36] In this context, characterizing intra-tumor heterogeneity would seem to be necessary to better predict the clinical outcome of a specific tumor at the moment of diagnosis and to establish the most appropriate treatment. The standard treatment for endometrial cancer is well established, involving surgery followed by adjuvant radiotherapy in tumors with a high-risk of recurrence. Chemotherapy is usually restricted to metastatic/recurrent and high-grade endometrial cancers, although traditional chemotherapy regimens are less effective than in cancers of other organs.[5] In this sense, numerous clinical trials have been stratified according to genetic features, based on *PTEN, PIK3CA*, or *FGFR3* mutational status. Consequently, tumor heterogeneity represents a therapeutic challenge and the use of a single diagnostic biopsy of a tumor may be insufficient, leading to the misclassification of a significant proportion of patients.

Several studies have centered on the feasibility of using liquid biopsies to analyze intra-tumor genetic heterogeneity.[37–40] In endometrial cancer, uterine aspirates are used as minimally invasive and highly sensitive biopsies for histological diagnosis or molecular characterization.[16–18] In this regard, we found that the genetic analysis of uterine aspirates coupled to their pathological classification could be a very sensitive approach to detect endometrial malignant neoplasia. This implies that detecting a cancer-related mutation (such as those detected by the method we employed) is related to a possible malignant disorder or tumor. Although this seems to be true in our series it remains controversial, and a significantly larger number of samples (both normal and malignant) should be analyzed to address this issue. Paired sequencing of uterine aspirates and hysterectomy specimens confirms the efficacy in revealing malignant disorders (endometrial tumors or atypical hyperplasia) in uterine aspirates. Only three samples (5.5%) of uterine aspirates from tumor cases did not show any of the surgical tumor sample mutations, whereas a total of 42 (77.8%) of the aspirates carried all the mutations found in the corresponding hysterectomy specimen.

Furthermore, we detected mutations in aspirates that could not be evaluated pathologically. The amount of tissue obtained from endometrial biopsies from postmenopausal patients is sometimes insufficient to obtain an adequate diagnosis, which in the majority of cases is due to the presence of endometrial atrophy. However, patients with endometrial cancer on occasions provided poor quality samples. In a recent study of 1120 endometrial samples classified as unsuitable for diagnosis, a second biopsy was obtained from 38% of the patients that was suitable for diagnosis in 75% of cases, with 10% having a malignant tumor.[23] Our results show that mutation analysis could indicate the presence of endometrial cancer or at least some pre-malignant anomaly, emphasizing the need for resampling in such cases and providing valuable information to accelerate the diagnosis.

Genetic analysis of uterine aspirates captures the intra-tumor heterogeneity identified in endometrioid endometrial carcinomas. The mutation discovery rate, defined as the percentage of mutations detected in each individual sample with respect to all the mutations found in a given patient, was used to measure the sensitivity of mutation detection in each sample. In heterogeneous tumors, the uterine

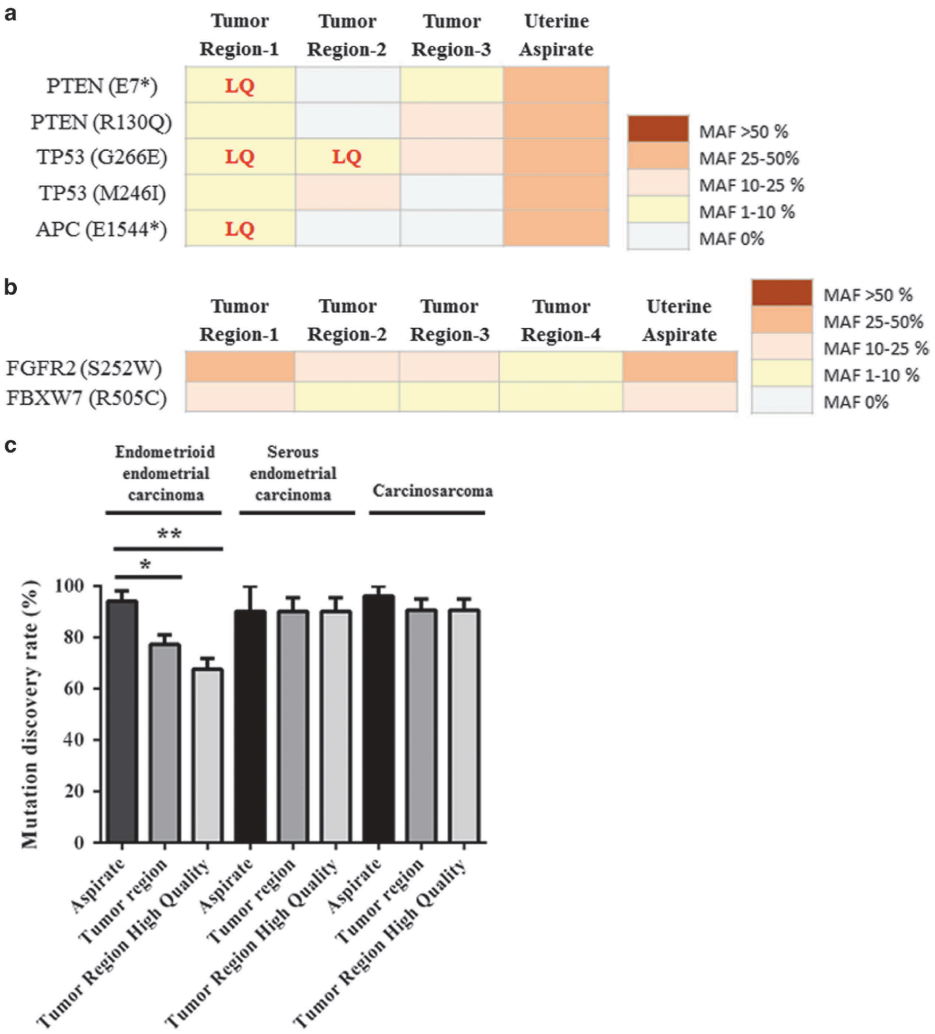


Figure 3 Characterization of the intra-tumor genetic heterogeneity in endometrial tumors. Representative mutational profile of genetically heterogeneous endometrioid carcinoma (**a**, EEC-1) and of a homogeneous endometrioid tumor (**b**, EEC-11). The colors in the squares represent the mutant allelic frequencies (MAFs). The squares marked as LQ identify low quality variants in the ion PGM analysis. The mutation discovery rate is defined as the percentage of mutations detected in each sample with respect to the totality of the mutations observed in all the samples analyzed from the same patient (see 'Materials and methods' section). The graph represents the mean mutation discovery rate (**c**) in endometrioid carcinomas, serous carcinomas, and carcinosarcomas. (*$0.005 < P < 0.05$; **$0.001 < P < 0.005$).

aspirate mutation discovery rate was higher than that in at least one of the tumor regions. In fact, the mutation discovery rate value was higher in the uterine aspirate than in the tumor regions used for pathological diagnosis (tumor region 1) in 8 of the 10 heterogeneous endometrioid carcinomas. These results highlight the potential utility of this type of biopsy and reveal that the use of a unique tumor sample in diagnosis could underestimate the mutational burden in heterogeneous tumors. However, the study of multiple samples of a given tumor as a routine practice is still a difficult issue, as it would increase significantly the time and cost of diagnosis. Moreover, combining DNA from different tumor samples previously to the targeted sequencing is not a good option, because it would lead to a decrease in the frequency of those mutations, which are not present in all the tumor regions, causing some low-requency variants to be undetected. It is also worth pointing that it is fairly difficult to calculate how many tumor regions need to be analyzed to cover the intra-tumor heterogeneity found in each case. Taken together, these arguments increase the value of uterine aspirates as a genetic

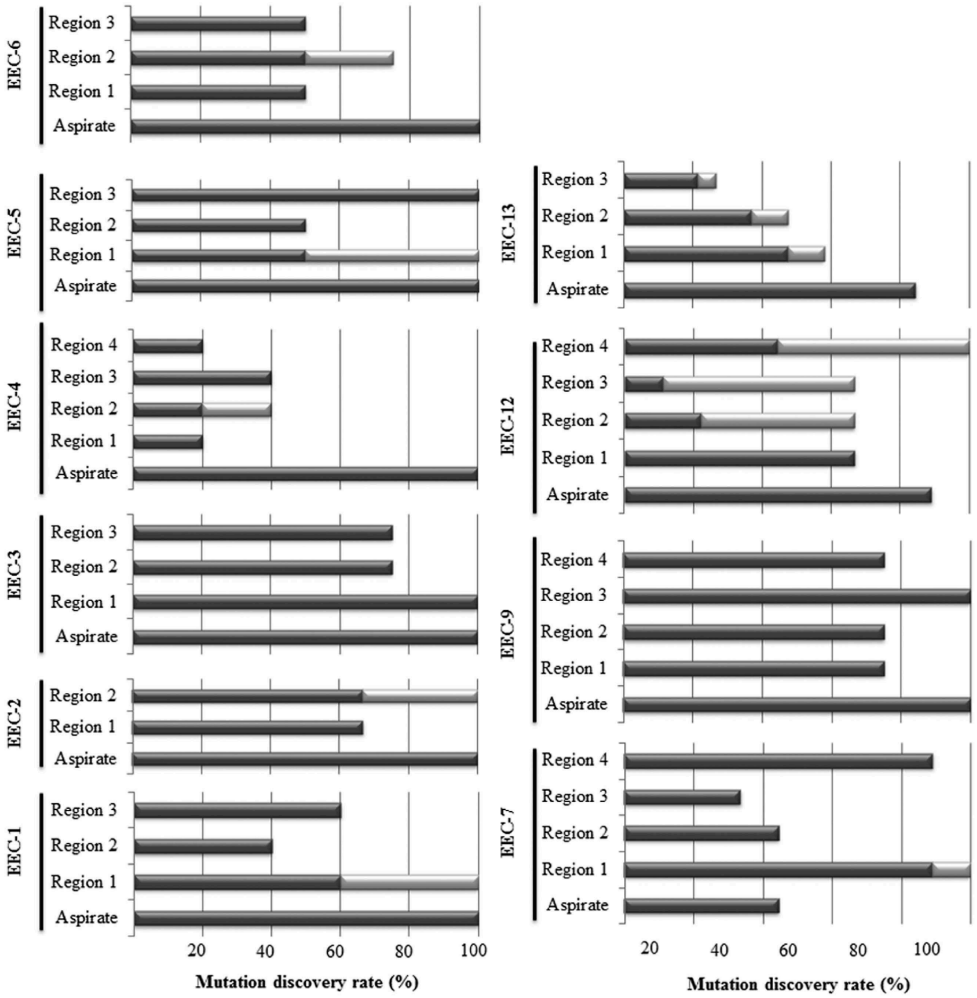


Figure 4 Mutation discovery rate in heterogeneous endometrioid endometrial carcinoma. The mutation discovery rate was calculated for each sample from the heterogeneous endometrioid carcinoma patients as indicated in the 'Materials and methods' section. Each bar represents a sample, from the bottom to the top: uterine aspirate and the different tumor regions.[1–4] The dark gray color represents the percentage of high quality variants detected and the light gray reflects the LQ variants identified in the ion PGM analysis.

diagnostic biopsy, solving, at least in part, some of the problems found in the study of hysterectomy specimens. The fact that intra-tumor heterogeneity may be represented in uterine aspirates is probably related to the nature of such samples, consisting of cells from many different parts of the uterine cavity, which could provide a more representative picture of the entire tumor specimen than samples from a specific tumor region. Similar results were observed in ovarian carcinomas where intra-tumor genetic heterogeneity was evident when solid tumor biopsies were compared,[32–34] but not when different ascites from the same patient were compared.[41] In this case, ascites could represent the entire cavity in a similar way that uterine aspirates do in uterine cancers, capturing all the genetic mutations and representing the heterogeneity found in the solid tumor biopsies.

The use of non-invasive biopsies to diagnose and characterize tumors is currently a relevant clinical challenge. The data presented here shed light on the molecular characterization of minimally-invasive biopsies in endometrial cancer, and they provide potential solutions to the problem of detecting genetic heterogeneity, as well as valuable information in the case of biopsies with insufficient material. These data pave the way for the use of such analyses for other diseases.

Acknowledgments

We thank all those at the Translational Research Laboratory and Immunohistochemical Laboratory from MD Anderson Madrid for their invaluable help. Tissue samples were obtained with the support of MD Anderson Foundation Biobank (Record Number B.0000745, ISCIII National Biobank Record), the 'Xarxa Catalana de Bancs de Tumors' and 'Plataforma de Biobancos' ISCIII (PT13/0010/0014, B.000609). This work was supported by grants from the AECC (Grupos Estables de Investigacion 2011-AECC- GCB 110333 REVE), the 'Fundació La Marató, TV3' (2/C/2013) to AG-M, JR, XM-G, and GM-B; Instituto de Salud Carlos III (ISCIII) (PI13/00132 and RETIC-RD12/0036/0007 to GM-B; RETIC-RD12-/0036/0035 to JR; PI13/01701, and RD12/0036/0013 to XM-G; PI14/02043 and PI14/01942 to AG and MA); the 'CIRIT, Generalitat de Catalunya' (2014 SGR 1330 to JR; and 2014 SGR 138 to XM-G), GEIS award 2013 to GM-B and PG-S, and the 'Communidad de Madrid' (S2010/BMD-2303) to GM-B. AM is funded by the Spanish Ministry of Education, Culture, and Sports (FPU2012-5338). IC and PG-S are funded by PhD and postdoctoral contracts, respectively, from the AECC Scientific Foundation. EC is funded by the Spanish Ministry of Economy and Competitiveness (FPDI-2013-18322).

Author contributions

AM and PG-S performed the sequencing experiment and analysis. AM, EC, PG-S, and IC contributed to the sample processing. AR-S, SG, BD-F, AV, and AG performed the pathological analysis of the samples. LC, SA, AG-M, XG-T, PZ-M, and MB helped obtain the samples. MA and RL-L read and corrected the manuscript. EC, XM-G, JR, and GM-B conceived the study, participated in its design, and helped draft the manuscript. GM-B discussed and directed the study. All the authors read and approved the final manuscript.

Disclosure/conflict of interest

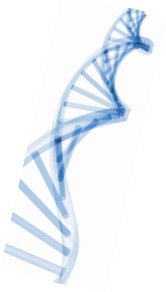
The authors declare no conflict of interest.

References

1 Ferlay J, Soerjomataram I, Dikshit R, *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–E386.
2 Bokhman JV. Two pathogenetic types of endometrial carcinoma. Gynecol Oncol 1983;15:10–17.
3 Lax SF, Kurman RJ. A dualistic model for endometrial carcinogenesis based on immunohistochemical and molecular genetic analyses. Verh Dtsch Ges Pathol 1997;81:228–232.
4 Llaurado M, Ruiz A, Majem B, *et al.* Molecular bases of endometrial cancer: new roles for new actors in the diagnosis and the therapy of the disease. Mol Cell Endocrinol 2012;358:244–255.
5 Yeramian A, Moreno-Bueno G, Dolcet X, *et al.* Endometrial carcinoma: molecular alterations involved in tumor development and progression. Oncogene 2013;32:403–413.
6 Cancer Genome Atlas Research Network, Kandoth C, Schultz N, *et al.* Integrated genomic characterization of endometrial carcinoma. Nature 2013;497:67–73.
7 Hong B, Le Gallo M, Bell DW. The mutational landscape of endometrial cancer. Curr Opin Genet Dev 2015;30C:25–31.
8 Zighelboim I, Goodfellow PJ, Gao F, *et al.* Microsatellite instability and epigenetic inactivation of MLH1 and outcome of patients with endometrial carcinomas of the endometrioid type. J Clin Oncol 2007;25:2042–2048.
9 Tritz D, Pieretti M, Turner S, *et al.* Loss of heterozygosity in usual and special variant carcinomas of the endometrium. Hum Pathol 1997;28:607–612.
10 Cantrell LA, Blank SV, Duska LR. Uterine carcinosarcoma: a review of the literature. Gynecol Oncol 2015;137:581–588.
11 Artioli G, Wabersich J, Ludwig K, *et al.* Rare uterine cancer: carcinosarcomas. Review from histology to treatment. Crit Rev Oncol Hematol 2015;94:98–104.
12 Colombo N, Preti E, Landoni F, *et al.* Endometrial cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2013;24: vi33–vi38.
13 Clark TJ, Mann CH, Shah N, *et al.* Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial cancer: a systematic quantitative review. BJOG 2002;109:313–321.
14 Karateke A, Tug N, Cam C, *et al.* Discrepancy of pre- and postoperative grades of patients with endometrial carcinoma. Eur J Gynaecol Oncol 2011;32:283–285.
15 Wang XY, Pan ZM, Chen XD, *et al.* Accuracy of tumor grade by preoperative curettage and associated clinicopathologic factors in clinical stage I endometriod adenocarcinoma. Chin Med J 2009;122:1843–1846.
16 Colas E, Perez C, Cabrera S, *et al.* Molecular markers of endometrial carcinoma detected in uterine aspirates. Int J Cancer 2011;129:2435–2444.
17 Perez-Sanchez C, Colas E, Cabrera S, *et al.* Molecular diagnosis of endometrial cancer from uterine aspirates. Int J Cancer 2013;133:2383–2391.
18 Stelloo E, Nout RA, Naves LC, *et al.* High concordance of molecular tumor alterations between pre-operative curettage and hysterectomy specimens in patients with endometrial carcinoma. Gynecol Oncol 2014;133: 197–204.
19 Jamal-Hanjani M, Quezada SA, Larkin J, *et al.* Translational implications of tumor heterogeneity. Clin Cancer Res 2015;21:1258–1266.
20 Sims D, Sudbery I, Ilott NE, *et al.* Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet 2014;15:121–132.
21 McCluggage WG. My approach to the interpretation of endometrial biopsies and curettings. J Clin Pathol 2006;59:801–812.
22 Phillips V, McCluggage WG. Results of a questionnaire regarding criteria for adequacy of endometrial biopsies. J Clin Pathol 2005;58:417–419.
23 Kandil D, Yang X, Stockl T, *et al.* Clinical outcomes of patients with insufficient sample from endometrial

biopsy or curettage. Int J Gynecol Pathol 2014;33:500–506.
24 Durrett R, Foo J, Leder K, *et al.* Intratumor heterogeneity in evolutionary models of tumor progression. Genetics 2011;188:461–477.
25 Gerlinger M, Rowan AJ, Horswell S, *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883–892.
26 Khalique L, Ayhan A, Whittaker JC, *et al.* The clonal evolution of metastases from primary serous epithelial ovarian cancers. Int J Cancer 2009;124:1579–1586.
27 Cooke SL, Ng CK, Melnyk N, *et al.* Genomic analysis of genetic heterogeneity and evolution in high-grade serous ovarian carcinoma. Oncogene 2010;29:4905–4913.
28 Lawrence MS, Stojanov P, Mermel CH, *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 2014;505:495–501.
29 Alexandrov LB, Nik-Zainal S, Wedge DC, *et al.* Signatures of mutational processes in human cancer. Nature 2013;500:415–421.
30 Vogelstein B, Papadopoulos N, Velculescu VE, *et al.* Cancer genome landscapes. Science 2013;339:1546–1558.
31 Kleppe M, Levine RL. Tumor heterogeneity confounds and illuminates: assessing the implications. Nat Med 2014;20:342–344.
32 Khalique L, Ayhan A, Weale ME, *et al.* Genetic intratumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours. J Pathol 2007;211:286–295.
33 Bashashati A, Ha G, Tone A, *et al.* Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. J Pathol 2013;231:21–34.
34 Mota A, Trivino JC, Rojo-Sebastian A, *et al.* Intra-tumor heterogeneity in TP53 null high grade serous ovarian carcinoma progression. BMC Cancer 2015;15:940.
35 Pribluda A, de la Cruz CC, Jackson EL. Intratumoral heterogeneity: from diversity comes resistance. Clin Cancer Res 2015;21:2916–2923.
36 Bhang HE, Ruddy DA, Krishnamurthy Radhakrishna V, *et al.* Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nat Med 2015;21:440–448.
37 Patel KM, Tsui DW. The translational potential of circulating tumour DNA in oncology. Clin Biochem 2015;48:957–961.
38 De Mattos-Arruda L, Weigelt B, Cortes J, *et al.* Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle. Ann Oncol 2014;25:1729–1735.
39 Raimondi C, Nicolazzo C, Gradilone A, *et al.* Circulating tumor cells: exploring intratumor heterogeneity of colorectal cancer. Cancer Biol Ther 2014;15:496–503.
40 Broersen LH, van Pelt GW, Tollenaar RA, *et al.* Clinical application of circulating tumor cells in breast cancer. Cell Oncol 2014;37:9–15.
41 Castellarin M, Milne K, Zeng T, *et al.* Clonal evolution of high-grade serous ovarian carcinoma from primary to recurrent disease. J Pathol 2013;229:515–524.

Supplementary Information accompanies the paper on *Modern Pathology* website (http://www.nature.com/modpathol)



ANNEX II: Supplementary Materials

Supplementary Table 1 FIGO and TNM classification of tumors of the ovary, fallopian tube and primary peritoneal carcinoma

FIGO Stages	TNM* Categories	Description
I	T1	Tumor limited to the ovaries
IA	T1a N0 M0	Tumor limited to one ovary (capsule intact,) or fallopian tube surface, no malignant cells in ascites or peritoneal washings
IB	T1b N0 M0	Tumor limited to one or both ovaries (capsule intact) or fallopian tubes; no tumor on ovarian or fallopian tube surface; no malignant cells in ascites or peritoneal washings
IC	T1c	Tumor limited to one or both ovaries or fallopian tubes with any of the following:
IC1	T1c1 N0 M0	Surgical spill
IC2	T1c2 N0 M0	Capsule ruptured before surgery or tumor on ovarian or fallopian tube surface
IC3	T1c 3N0 M0	Malignant cells in ascites or peritoneal washings
II	T2	Tumor involves one or both ovaries or fallopian tubes with pelvic extension below pelvic brim or primary peritoneal cancer
IIA	T2a N0 M0	Extension and/or implants on uterus and/or fallopian tubes and/or ovaries.
IIB	T2b N0 M0	Extension to other pelvic intraperitoneal location
III	T3 and/or N1	Tumor involves one or both ovaries or fallopian tubes or primary peritoneal carcinoma, with cytologically or histologically confirmed spread to the peritoneum outside the pelvis and/or metastasis to the retroperitoneal lymph nodes
IIIA1	T1/2 N1 M0	Retroperitoneal lymph node metastasis only
IIIA1i		Lymph node metastasis up to 10 mm in greatest dimension
IIIA1ii		Lymph node metastasis more than 10 mm in greatest dimension
IIIA2	T3a N0/N1 M0	Microscopic extrapelvic (above the pelvic brim) peritoneal involvement with or without retroperitoneal lymph node
IIIB	T3b N0/N1 M0	Macroscopic peritoneal metastasis beyond the pelvis up to 2 cm in greatest dimension with or withouth retroperitoneal lymph node metastasis
IIIC	T3c N0/N1 M0	Macroscopic peritoneal metastasis beyond the pelvis more than 2 cm in greatest dimension, with or without retroperitoneal lymph node metastasis (excludes extension of tumor to capsule of liver and spleen without parenchymal involvement of either organ)
IV	Any T Any N M1	Tumor involves one or both ovaries with distant metastasis excluding peritoneal metastasis
IVA		Pleural effusion with positive cytology
IVB		Parenchymal metastasis and metastasis to extra-abdominal organs (including inguinal lymph nodes and lymph nodes outside the abdominal cavity)

* T: primary tumor, N: regional lymph nodes**, M: distant metastasis (0: absence, 1: presence)

** The regional lymph nodes are the hypogastric (obturator), common iliac, external iliac, lateral sacral, para-aortic, and inguinal nodes.

Supplementary Table 2 FIGO and TNM classification of the uterine endometrium tumors

FIGO Stages	TNM* Categories	Description
	Tis	Carcinoma in situ (preinvasive carcinoma)
I[a]	T1	Tumor confined to the corpus uteri[a]
IA[a]	T1a N0 M0	Tumor limited to endometrium or invading less than half of myometrium
IB	T1b N0 M0	Tumor invades one half or more of myometrium
II	T2 N0 M0	Tumor invades cervical stroma, but does not extend beyond the uterus
III	T3 and/or N1	Local and/or regional spread as specified below
IIIA	T3a N0 M0	Tumor invades the serosa of the corpus uteri or adnexae (direct extension or metastasis)
IIIB	T3b N0 M0	Vaginal or parametrial involvement (direct extension or metastasis)
IIIC	T1, T2, T3 N1, N2 M0	Metastasis to pelvic or para-aortic lymph nodes[b]
IIIC1	T1, T2, T3 N1 M0	Metastasis to pelvic lymph nodes
IIIC2	T1, T2, T3 N2 M0	Metastasis to para-aortic lymph nodes with or without metastasis to pelvic lymph nodes
IVA	T4 Any N M0	Tumor invades bladder/bowel mucosa[c]
IVB	Any T Any N M1	Distant metastasis (excludes metastasis to vagina, pelvic serosa or adnexae)

a Endocervical glandular involvement only should now be considered as stage I. b Positive cytology has to be reported separately withouth changing the stage. C The presence of bullous oedema is not sufficient evidence to classify as T4. This lesion should be confirmed by biopsy.

* T: primary tumor, N: regional lymph nodes**, M: distant metastasis; (0: absence, 1/2: presence). Differentation between N1 (pelvic lymph nodes) and N2 (para-aortic lymph nodes) was applied due to significant differences in survival were found [287, 288].

Supplementary Table 3 Frequently mutated genes in endometrial carcinoma

Rank	Gene	Gene name	Discovery in Cancer
1	*PTEN*	phosphatase and tensin homolog	Known in EC
2	*PIK3CA*	phosphoinositide-3-kinase, catalytic, alpha polypeptide	Known in EC
3	*PIK3R1*	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)	Known in EC
4	*CTNNB1*	catenin beta 1	Known in EC
5	*TP53*	tumor protein p53	Known in EC
6	*KRAS*	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog	Known in EC
7	*FBXW7*	F-box and WD repeat domain containing 7	Known in EC
8	*SPOP*	speckle-type POZ protein	Known in EC
9	*CTCF*	CCCTC-binding factor	Known in EC
10	*ARID1A*	AT rich interactive domain 1A	Known in EC
11	*PPP2R1A*	protein phosphatase 2 (formerly 2A), regulatory subunit A , alpha isoform	Known in EC
12	*CCND1*	cyclin D1	Known in EC
13	*FGFR2*	fibroblast growth factor receptor 2	Known in EC
14	*ARHGAP35*	glucocorticoid receptor DNA binding factor 1	Known in EC
15	*CHD4*	chromodomain helicase DNA binding protein 4	Known in EC
16	*ZFHX3*	zinc finger homeobox 3	Known in EC
17	*SOX17*	SRY (sex determining region Y)-box 17	Known in EC
18	*ERBB3*	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3	Known in EC
19	*ARID5B*	AT rich interactive domain 5B	Known in EC
20	*NRAS*	neuroblastoma RAS viral (v-ras) oncogene homolog	Known in EC
21	*NFE2L2*	nuclear factor (erythroid-derived 2)-like 2	New in EC, but reported in other cancers
22	*MAX*	MYC associated factor X	Novel SMG
23	*SOS1*	son of sevenless homolog 1 (Drosophila)	Known in EC
24	*SGK1*	serum/glucocorticoid regulated kinase 1	Known in EC
25	*ERBB2*	v v-erb-b2 erythroblastic leukemia viral oncogene homologue 2	New in EC, but reported in other cancers
26	*BCOR*	BCL6 co-repressor	Known in EC
27	*ESR1*	estrogen receptor 1	Novel SMG
28	*RRAS2*	related RAS viral (r-ras) oncogene homolog 2	Novel SMG
29	*SIN3A*	SIN3 homolog A, transcription regulator (yeast)	Novel SMG
30	*MYCN*	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived	Known in EC
31	*AFMID*	arylformamidase	Novel SMG
32	*MECOM*	MDS1 and EVI1 complex locus	Novel SMG
33	*FOXA2*	forkhead box A2	Novel SMG
34	*ALPK2*	alpha-kinase 2	New in EC, but reported in other cancers
35	*AKT1*	v-akt murine thymoma viral oncogene homolog 1	Known in EC
36	*METTL14*	methyltransferase like 14	Novel SMG
37	*SERHL2*	serine hydrolase-like 2	Novel SMG
38	*WDR45*	WD repeat domain 45	Novel SMG
39	*U2AF1*	U2 small nuclear RNA auxiliary factor 1	New in EC, but reported in other cancers
40	*TAB3*	TGF-beta activated kinase 1/MAP3K7 binding protein 3	Novel SMG
41	*ADNP*	activity-dependent neuroprotector homeobox	Known in EC
42	*MARK3*	MAP/microtubule affinity-regulating kinase 3	Novel SMG
43	*EDC4*	enhancer of mRNA decapping 4	Novel SMG
44	*DICER1*	dicer 1, ribonuclease type III	Novel SMG
45	*RBM39*	RNA binding motif protein 39	Novel SMG
46	*POLE*	polymerase (DNA directed), epsilon	Known in EC
47	*RNF43*	ring finger protein 43	Novel SMG
48	*JAK1*	Janus kinase 1 (a protein tyrosine kinase)	Novel SMG
49	*NRIP1*	nuclear receptor interacting protein 1	Novel SMG

Adapted from Gibson *et al.*, 2016[50].

Supplementary Table 4 Primary antibodies used in immunohistochemistry analysis

Protein	Antibody	Trading House	Dilution	Staining pattern
WT1	Monoclonal Mouse AntHuman Wilms' Tumor 1 (WT1) Protein Clon 6F-H2i-	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
PTEN	Monoclonal Mouse Anti-Human PTEN Clone 6H2.1	Dako (Glostrup, Denmark)	1:100	Nuclear and cytoplasmic
p53	Monoclonal Mouse Anti-Human p53 Protein Clone DO-7	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
Ki-67	Monoclonal Mouse Anti-Human Ki-67 Antigen Clone MIB-1	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
MLH1	Monoclonal Mouse Anti-Human MLH1 Clone ES05	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
MSH2	Monoclonal Mouse Anti-Human MSH2 Clone E11	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
MSH6	Monoclonal Mouse Anti-Human MSH6 Clone EP49	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
PMS2	Policlonal Rabbit Anti-Human PMS2 Clone EP51	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
PR	Monoclonal Mouse Anti-Human R.P. Antigen Clone PgR636	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
ER	Monoclonal Mouse Anti-Human R.E. Antigen Clone 1D5	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
HMGA2	Polyclonal Rabbit Anti-Human HMGA2	LifeSpan BioSciences (Seatle, WA)	1/100	Nuclear
IMP2	Monoclonal Rabbit Anti- IGF2BP2 Clone EPR6741 (B)	Abcam (Cambridge, MA)	1/100	Cytoplasmic
IMP3	Monoclonal Mouse Anti-Human IMP3 Clone 69.1	Dako (Glostrup, Denmark)	1/100	Cytoplasmic
Cyclin E1	Polyclonal Rabbit Anti-Human Cyclin E1	Sigma-Aldrich (St.Louis, MO)	1/100	Nuclear
p16	CINtec Histology Kit	Dako (Glostrup, Denmark)	Ready-to-use	Nuclear
FolR1	Polyclonal Rabbit Anti-Human Folate Binding Protein	Abcam (Cambridge, MA)	1/100	Membranous
CK7	Monoclonal Mouse Anti-Human CK7 Antigen Clone OV-TL	Dako (Glostrup, Denmark)	Ready-to-use	Cytoplasmic
Beta-catenin	Monoclonal Mouse Anti-Human beta-cat Antigen Clone b-catenin-1	Dako (Glostrup, Denmark)	Ready-to-use	Cytoplasmic and membranous / Nuclear (mutated)

Supplementary Table 5 Ampliseq custom panel designs

Supplementary Table 5 is included in the extra CD information.

Supplementary Table 6 PCR conditions for Sanger analysis of variants detected by NGS in the ovarian cancer project (A), the uterine aspirate project (B) and the analysis of microsatellite instability (C)

A) Ovarian cancer project

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Length	Annealing temperature
CNOT1	TGCGTTGTTTTTGTCTTGCT	AGCCAGACCTAGTGCCATGT	225	58
CSMD3	TGAATGAGCCCTTTTGTTTTT	GGCAGTTTTACCCAACCACT	180	58
CTC1	TCAAAGGAAACACTGGCACA	CCCTTGCTCTTGGTCTTTCTT	162	58
ECE1	ACACGAATTCCCTCTCATGC	CCTGTTTCACCCATCAGGTC	234	58
FAP	TTCTCCAGCTCCCTTCAGTC	CGTGTTAAATGCTTTCACAGTAACA	250	58
FRMPD1	CAGAAGGCAAAAGTGACAGC	CCTGCAAAACCCAAAGATGT	127	58
GPNMB	TCAGAAGCAAAGGCCTGAGT	CAACTTCCCCAAACCACAAT	235	58
GTF2I	GTGTGATCCAGAGCTGCAAA	AGGTGTGGGAGTTAAACAGCA	214	58
HEPHL1	GGCCTGTGTTTTGCCTTTAG	CTGGTTTTGTCATGGGCACT	161	58
KIF21A	ATCACAGCATTCAGTTTACAACC	TCTGCCCTTGTTCATATATCCAT	176	58
KIF21B	TCACAACAGCCTTTCTGCAC	ACTGATTTGCTGGCTGGTCT	168	58
KMT2A	GAGGAAGACCTCCCACCTTC	TTTGTACCCCCTTCCTTCCT	231	58
LAMA2	CCAGTGCATAGGCATGTACC	GCACTTGGTCTCCCATTGAT	176	58
OR56B1	CATTCACAGCTGGCAACACT	GCCAGACCCATGTCTACCAT	165	58
PCYT2	CAGAAGAAGGAAGCCAAGGA	CAAGGAGGCAGAGTCCTCAC	223	58
PLXNA1	CGTACTGGAGCCACTCAGC	CGATGAGCACCGTGTAGTTG	212	58
ROBO2	TACAAAGATGGGGAGCGAGT	CCAAGATAGTTCCTCGCAACA	209	58
SMG7	TTTAGCAATGAAACCGAGCA	GCTTTGCTACATCGATGAAATG	153	58
TFDP1	AGATGTCCAGGCCAACTCCT	CTTCTTGCTGGTGTTGACGA	241	58
TP53	CCACAGGTCTCCCCAAGG	TGGCAAGTGGCTCCTGAC	183	55
TRERF1	TCCACATCCTTGATGGGTTT	ACAGAGGCAAAAGGCTCAGA	138	58
UBE2D4	CACCAGGAGAATTCCTTCCA	ACTTACAGTCATCACCGACAGG	162	58
UBR3	TTGGACAGAATATCGGGCTTA	TGGACATACGTCATGGCTTG	202	58
ZFAT	CAGCAGGTGTCTCAGGTCAA	GCTGCCTTTCCTTACCACAG	248	58
ZNF664	ACACCTCCAGCCTCTGCAT	AGGCCTTTCCGCACTCAT	246	58

B) Uterine aspirate project

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Length	Annealing temperature
ABL1	ATCACCACGCTCCATTATCC	CAGTCCCAGCCTACCTTCAA	209	55
APC	TCCAGGTTCTTCCAGATGCT	TTTCCTGAACTGGAGGCATT	156	57
BRAF	TCATAATGCTTGCTCTGATAGGA	GGCCAAAATTTAATCAGTGGA	224	52
CDKN2A	CAGCTCCTCAGCCAGGTC	CCTGGCTCTGACCATTCTGTT	245	57
CTNNB1	ATGGAACCAGACAGAAAAGC	GGTACTTGTTCTTGAGTGAAG	200	55
FGFR2	GCTGCCCATGAGTTAGAGGA	TATTTGGGCGAATGCAGTTT	201	57
GNAS	TGGCTTTGGTGAGATCCATT	GGACTGGGGTGAATGTCAAG	176	55
KRAS	GTGTGACATGTTCTAATATAGTCA	GAATGGTCCTGCACCAGTAA	214	56
PIK3CA	CCGTGAGGCTACATTAATAACC	AAGCTTTATGGTTATTTGCATTTT	288	55
PIK3CA	ATTCTCAACTGCCAATGGA	AGTGCAAGAAAAAGGTTATCTAAAA	244	55
PIK3CA	TGAATTTTCCTTTTGGGGAAG	GAGAGAAGGTTTGACTGCCATAA	292	57
PIK3CA	TTGAAAATGTATTTGCTTTTTCTGT	AACAGAGAATCTCCATTTTAGCA	249	57
PTEN	TTGACCAATGGCTAAGTGAAGA	TCTCAGATCCAGGAAGAGGAA	217	57
PTEN	TTTTTCAATTTGGCTTCTCTTTTT	TGTTCCAATACATGGAAGGATG	220	55
PTEN	AAAGGCATTTCCTGTGAAATAA	TTGGATATTTCTCCCAATGAAAG	249	55
PTEN	ACCAGGACCAGAGGAAACCT	AGTCAACAACCCCCACAAAA	224	55
SMAD4	TGGGAAGAGATCACCCTGTC	GGCCCGGTGTAAGTGAATTT	200	55
TP53	CCGTGTTCCAGTTGCTTTATC	AGCCCTGTCGTCTCTCCA	288	58
TP53	CCACAGGTCTCCCCAAGG	TGGCAAGTGGCTCCTGAC	183	55
TP53	CCTATCCTGAGTAGTGGTAA	TCCTCCACCGCTTCTTGT	192	55

C) Microsatellite instability analysis

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Length	Annealing temperature
BAT25	TCGCCTCCAAGAATGTAAGT	TCTGGATTTTAACTATGGCTC	100-150	55
BAT26	TGACTACTTTTGACTTCAGCC	AACCATTCAACATTTTTAACC	100-150	55
D2S123	AAACAGGATGCCTGCCTTTA	GGACTTTCCACCTATGGGAC	150-300	55
D5S346	ACTCACTCTAGTGATAAATCGGG	AGCAGATAAGACAAGTATTACTAG	100-150	55
D17S250	GGAAGAATCAAATAGACAAT	GCTGGCCATATATATATTTAAACC	100-200	55

Supplementary Table 7 Genetic variants idenfied by WES in *TP53* null HGSOC samples

Supplementary Table 7 is included in the extra CD information.

Supplementary Table 8 Functional annotation of the genetic variants identified by WES

Category	Term	p-Value	Genes
BIOCARTA	**Cyclins and Cell Cycle Regulation**	0.008	*CDK7, CDC25A, TFDP1*
BIOCARTA	**Cell Cycle: G1/S Check Point**	0.010	*TP53, CDC25A, TFDP1*
COG ONTOLOGY	**Cell division and chromosome partitioning**	0.194	*MYH3, CIT, PLEC*
GOTERM_BP_ALL	GO:0007155: **Cell adhesion**	0.003	*LAMA2, PCDHA7, MYBPC2, DSG2, FAT3, TRO, ZAN, ROBO2, MSLNL, GPNMB, SIRPA*
GOTERM_BP_ALL	GO:0022610: **Biological adhesion**	0.003	*LAMA2, PCDHA7, MYBPC2, DSG2, FAT3, TRO, ZAN, ROBO2, MSLNL, GPNMB, SIRPA*
GOTERM_BP_ALL	GO:0007156: **Homophilic cell adhesion**	0.004	*PCDHA7, DSG2, FAT3, TRO, ROBO2*
GOTERM_BP_ALL	GO:0016337: **Cell-cell adhesion**	0.013	*PCDHA7, DSG2, FAT3, TRO, ZAN, ROBO2*
GOTERM_BP_ALL	GO:0000059: **Protein import into nucleus, docking**	0.083	*IPO8, KPNB1*
GOTERM_BP_ALL	GO:0007018: **Microtubule-based movement**	0.113	*KIF21A, KIF21B, DNAH5*
GOTERM_BP_ALL	GO:0006629: **Lipid metabolic process**	0.120	*ACADVL, ST6GALNAC6, TAZ, INSIG1, ETFDH, SULT2B1, PCYT2, TRERF1*
GOTERM_BP_ALL	GO:0006635: **Fatty acid beta-oxidation**	0.134	*ACADVL, ETFDH*
GOTERM_BP_ALL	GO:0009062: **Fatty acid catabolic process**	0.168	*ACADVL, ETFDH*
GOTERM_BP_ALL	GO:0019395: **Fatty acid oxidation**	0.181	*ACADVL, ETFDH*
GOTERM_BP_ALL	GO:0034440: **Lipid oxidation**	0.181	*ACADVL, ETFDH*
GOTERM_BP_ALL	GO:0008283: **Cell proliferation**	0.182	*INSIG1, TP53, CDK7, CDC25A, TFDP1*
GOTERM_BP_ALL	GO:0006351: **Transcription, DNA-dependent**	0.187	*GTF2I, TP53, CDK7, TTF2*
GOTERM_BP_ALL	GO:0006355: **Regulation of transcription, DNA-dependent**	0.187	*ZNF519, ZNF568, TP53, ZNF813, ZNF792, CDK7, TRERF1, NR2C1, BARX1, TRAK2, ZNF71, NOBOX, TFDP1*
GOTERM_MF_FAT	GO:0003774: **Motor activity**	0.008	*MYH15, MYH3, KIF21A, KIF21B, DNAH5*
GOTERM_MF_FAT	GO:0003777: **Microtubule motor activity**	0.068	*KIF21A, KIF21B, DNAH5*
GOTERM_MF_FAT	GO:0004386: **Helicase activity**	0.182	*DDX24, TTF2, SETX*
KEGG PATHWAY	hsa04110: **Cell cycle**	0.020	*TP53, CDK7, CDC25A, TFDP1*
REACTOME PATHWAY	REACT_578: **Apoptosis**	0.109	*DSG2, KPNB1, PLEC*

Supplementary Table 9 Targeted massive parallel sequencing results obtained in the uterine aspirates project

A) Paired samples of uterine aspirates (aspirate) and surgical specimens (hysterectomy) from atypical hyperplasia (AH), endometrioid endometrial carcinomas (EEC), serous endometrial caricnomas (SEC) and carcinosarcoma patients (CS). B) Additional tumor regions analyzed in heterogeneity study.

Supplementary Table 9 is included in the extra CD information.

Supplementary Table 10 Quality metrics of Ion PGM sequencing in the uterine aspirates project

A)

	Mean
Mapped Reads	236535.2
Bases	27070288.7
>=20QBases	25257514.2
>20QBases/Bases*100	93.3
On Target (%)	95.5
Mean Depth	1013.8
Uniformity (%)	97.6

B)

	Mean>=20QBases/ Bases*100
FFPE Aspirate	93.50
Frozen Aspirate	93.03
FFPE Tumor	93.27

C)

	pValue
Mapped Reads	0.100
Bases	0.109
>=20QBases	0.108
>=20QBases/Bases*100	0.712
On Target (%)	0.887
Mean Depth	0.133
Uniformity (%)	0.259

A) Quality metrics mean of Ion PGM sequencing for all the analyzed samples. Mapped reads refer to the number of reads that were aligned to the full reference genome. Bases are the post filtering base output. The parameter >=20Q Bases value represents the number of called bases with a predicted quality greater than or equal to 20Q, the predicted quality values being reported on the Phred scale, defined as -10log10 (error probability). Therefore, 20Q corresponds to a predicted error rate of one percent. On Target value is the percentage of mapped reads that were aligned over a target region, while the Mean Depth is the average base coverage depth over all the bases targeted in the reference genome. Uniformity is the percentage of target bases covered by at least 0.2X the average base read depth. B) Quality sequencing comparison between FFPE, frozen aspirates and FFPE tumors, using the proportion of >=20Q Bases and the total bases for each sample as a standard quality criteria. C) Quality metrics compared between paired samples of uterine aspirates and primary tumors with a paired T-test.

Supplementary Table 11 Summary of the histological grade an mutational profile identified in endometrial cancers and their paired uterine aspirates

Patient	Grade	Common variants (Aspirate and Hysterectomy)[a]	Hysterectomy variants (not detected in Aspirate)[b]	Aspirate variants (not detected in Hysterectomy)	% Hysterectomy variants detected in Aspirate[(a/(a+b))]
EEC-1	3	*PTEN* (2), *TP53* (2), *APC*	-	-	100
EEC-2	3	*PTEN, CTNNB1, CDKN2A*	-	-	100
EEC-3	3	*PTEN, PIK3CA* (2), *ABL1*	-	-	100
EEC-4	1	*PTEN, KRAS*	-	*PIK3CA* (3)	100
EEC-5	1	*PTEN, APC*	-	-	100
EEC-6	3	*FGFR2, PIK3CA, KIT*	-	*PIK3CA*	100
EEC-7	3	*FGFR2, PTEN* (2), *TP53*	*PTEN, TP53, SMARCB1, CTNNB1, CKN2A*	-	44
EEC-8	3	*PIK3CA, CTNNB1*	-	-	100
EEC-9	3	*FGFR2, PTEN, PIK3CA, CTNNB1*	-	-	100
EEC-10	3	*PIK3CA*	-	-	100
EEC-11	1	*FGFR2, FBXW7*	-	-	100
EEC-12	2	*PTEN* (3), *KRAS, RB1, ERBB2, TP53, PIK3CA, CTNNB1, FBXW7*	-	-	100
EEC-13	3	*NRAS, PTEN* (2), *ATM, HNF1A, PIK3CA, SMO, ABL1, CDKN2A*	*KRAS, GNA11* (2)	*PTEN, ATM, TP53* (2), *SMAD4, GNAS, CTNNB1*	75
EEC-14	3	*PTEN* (2), *ERBB2*	-	-	100
EEC-15	1	*PTEN*	*RET, STK11, PIK3CA*	*VHL*	25
EEC-16	3	*PTEN*(3), *TP53*	-	-	100
EEC-17	3	*KRAS, AKT1*	-	-	100
EEC-18	1	*PTEN, PIK3CA*	-	-	100
EEC-19	3	-	*FGFR2, PTEN* (3), *KRAS*	-	0
EEC-20	3	*KRAS*	-	-	100
EEC-21	1	-	-	*TP53, PIK3CA*	-
EEC-22	3	*PTEN* (2)	-	-	100
EEC-23	3	*FGFR2*	-	-	100
EEC-24	3	*PTEN, PIK3CA*	-	*CTNNB1, CDKN2A*	100
EEC-25	3	*TP53*	-	-	100
EEC-26	3	*KRAS, TP53*	-	-	100
EEC-27	3	*FGFR2, MLH1*	*PTEN*	-	67
EEC-28	2	*PTEN* (2), *CTNNB1*	-	-	100
EEC-29	2	*PTEN*	-	-	100
EEC-30	3	*PTEN, TP53, PIK3CA*	-	-	100
EEC-31	2	*PTEN*	-	-	100
EEC-32	2	*PTEN, CTNNB1*	-	-	100
EEC-33	1	*FGFR2, AKT1, CTNNB1*	-	-	100
EEC-34	2	*FGFR2, PTEN*	-	-	100
EEC-35	3	*PTEN* (3), *KRAS, PIK3CA*	*ATM, RB1, TP53, MET*	*RB1, KIT*	56
EEC-36	3	*PTEN, BRAF*	-	-	100
EEC-37	3	*PTEN, PIK3CA* (2)	-	-	100
EEC-38	3	*CTNNB1*	*TP53, VHL*	-	33
EEC-39	2	-	-	-	-
EEC-40	1	*FBXW7, PTEN*	-	-	100
EEC-41	3	*PTEN, FGFR2*	*FGFR2*	-	67
EEC-42	1	*PTEN*	*PIK3CA*	-	50
EEC-43	2	*PIK3CA, PTEN*	-	*TP53*	100
EEC-44	1	*PTEN, CTNNB1*	-	*KRAS*	100
SEC-1	3	*TP53*	-	-	100
SEC-2	3	*TP53*	-	-	100
SEC-3	3	*KRAS*	-	-	100
SEC-4	3	-	*ABL1*	*TP53*	0
SEC-5	3	*TP53, PIK3CA, BRAF, ATM*	-	-	100
SEC-6	3	-	*TP53*	-	0
SEC-7	3	*PIK3CA*	-	-	100

SEC-8	3	-	*TP53 (2), FBXW7*	-	0
SEC-9	3	*TP53*	-	-	100
CS-1	3	*PTEN, TP53, PIK3CA*	*FBXW7*	*KRAS*	75
CS-2	3	*TP53*	-	-	100
CS-3	3	*TP53*	-	-	100
CS-4	3	*TP53*	-	-	100
CS-5	3	*TP53*	-	-	100
CS-6	3	*PTEN, KRAS, PIK3CA*	-	-	100
CS-7	3	*KRAS, TP53*	*IDH2, TP53 (3), EGFR*	-	29
CS-8	3	*TP53*	-	-	100
CS-9	3	*FGFR2, PTEN*	-	-	100

- : mutations were not found

Supplementary Table 12 Summary of the mutational profile identified in atypical hyperplasia paired samples of hysterectomy specimens and uterine aspirates

Patient	Common variants (Aspirate and Hysterectomy)	Hysterectomy variants (not detected in Aspirate)	Aspirate variants (not detected in Hysterectomy)
AH-1	*PTEN*	-	*PTEN*
AH-2	-	-	*CTNNB1, PIK3CA, CDKN2A*
AH-3	-	-	*CTNNB1, PIK3CA*
AH-4	-	-	-
AH-5	-	-	-
AH-6	-	-	-
AH-7	*PIK3CA*	-	-
AH-8	-	-	-
AH-9	-	-	-
AH-10	-	-	*PIK3CA, PTEN, KRAS*

- : mutations were not found

Supplementary Table 13 Sanger sequencing study to determine *POLE* mutational status in the metastic endometrial cancer WES project

	Hotspot P286R	HotspotV411L
Forward primer (5' to 3')	CCCATCCCAGGAGCTTACTT	ATGTCCTCCGGGTCTAGCTC
Reverse primer (5' to 3')	GTGTTCAGGGAGGCCTAATG	TTGCATCTGTCTGTGTGGTG
Length	211	203
Annealing temperature (°C)	58	58
Start position	g.133,253,300 c.802-62	g.133,250,390 c.1227-97
End position	g.133,253,096 c.909+36	g.133,250,189 c.1331
Aminoacids covered	268-303	410-437

Supplementary Table 14 Somatic variants identified in the metastatic endometrial cancer WES project

Supplementary Table 14 is included in the extra CD information.

Supplementary Table 15 Somatic copy number alterations identified in the metastatic endometrial cancer WES project

Supplementary Table 15 is included in the extra CD information.

Supplementary Table 16 Mutational signatures identified in the metastatic endometrial cancer WES project

Sample	Mutational Signature
EEC-1_T1	Signature 1
EEC-1_T2	Signature 1
EEC-1_M	Signature 1
EEC-2_T1	Signature 1
EEC-2_T2	Signature 1
EEC-2_M	Signature 1
EEC-3_T1	Signature 1
EEC-3_T2	Signature 1
EEC-3_M	Signature 1
EEC-4_T1	Signature 1
EEC-4_T2	Signature 1
EEC-4_M	Signature 6*
EEC-5_T1	Signature 6
EEC-5_T2	Signature 6
EEC-5_M	Signature 6
EEC-6_T1	Signature 6
EEC-6_T2	Signature 6
EEC-6_M	Signature 6
EEC-7_T1	Signature 20
EEC-7_T2	Signature 20
EEC-7_M	Signature 6

Sample	Mutational Signature
SEC-1_T1	Signature 1
SEC-1_T2	Signature 1
SEC-1_T3	Signature 1
SEC-1_T4	Signature 1
SEC-1_M	Signature 1
SEC-2_T1	Signature 6*
SEC-2_T2	Signature 6*
SEC-2_T3	Signature 6*
SEC-2_M	Signature 6*
SEC-3_T1	Signature 1
SEC-3_T2	Signature 1
SEC-3_M1	Signature 6*
SEC-3_M2	Signature 6*
AEC_M	Signature 13
AEC_T2	Signature 13
AEC_T1	Signature 13

* Signature does not have correct associated indels. Patients with discordant signatures between its multiple samples are marked in red.

Supplementary Table 17 Quality metrics of Ion PGM sequencing in the validation of the metastatic endometrial cancer WES project

A)

	Mean
Mapped Reads	224512.3
Bases	28474526.6
>=20QBases	26424517.3
>20QBases/Bases*100	93.5
On Target (%)	93.8
Mean Depth	1242.6
Uniformity (%)	97.7

B)

	Mean>=20QBases/Bases*100
Frozen Tumor	93.2
FFPE Tumor	93.8

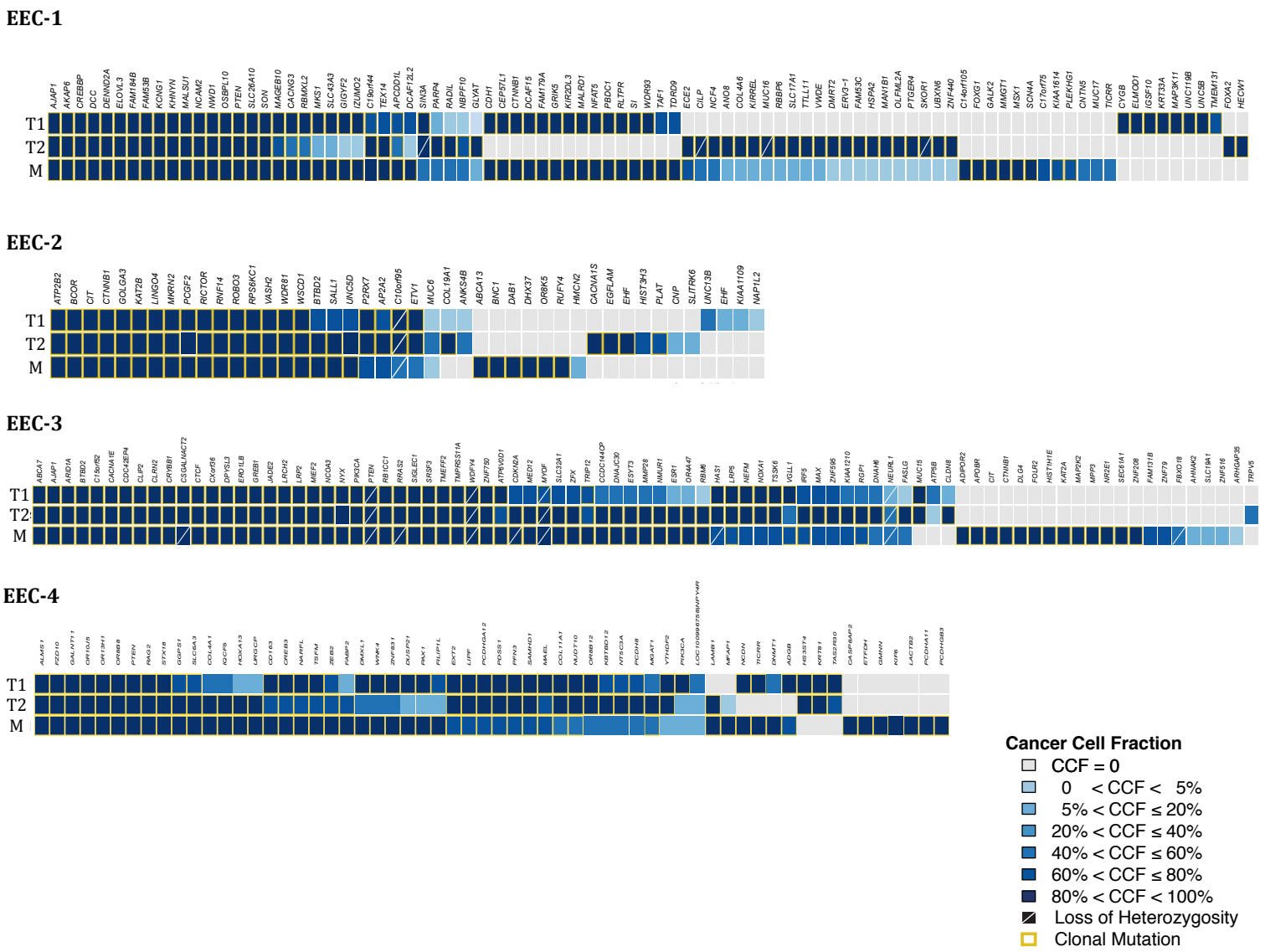
A) Quality metrics mean of Ion PGM sequencing for all the samples analyzed. Mapped reads refer to the number of reads that were aligned to the full reference genome. Bases are the post filtering base output. The parameter >=20Q Bases value represents the number of called bases with a predicted quality greater than or equal to 20Q, the predicted quality values being reported on the Phred scale, defined as -10log10 (error probability). Therefore, 20Q corresponds to a predicted error rate of one percent. On Target value is the percentage of mapped reads that were aligned over a target region, while the Mean Depth is the average base coverage depth over all the bases targeted in the reference genome. Uniformity is the percentage of target bases covered by at least 0.2X the average base read depth. B) Quality sequencing comparison between frozen and FFPE tumors, using the proportion of >=20Q Bases and the total bases for each sample as a standard quality criteria.

Supplementary Table 18 Targeted sequencing results obtained in the validation of the metastatic endometrial cancer WES project

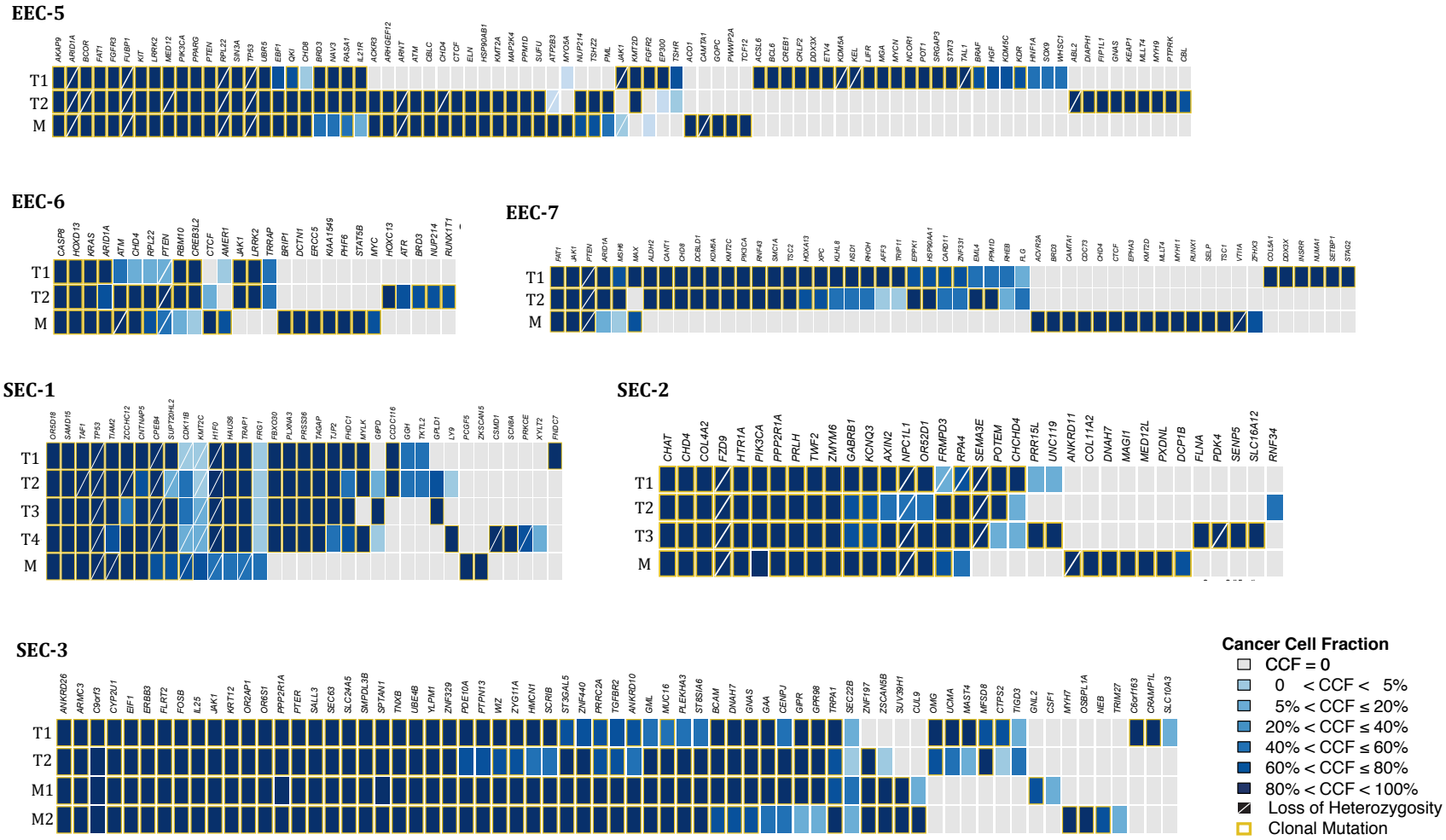
Supplementary Table 18 is included in the extra CD

Supplementary Table 19 Somatic variants identified by WES in the PDX tumors derived from an ambiguous endometrial carcinoma

Supplementary Table 19 is included in the extra CD



Supplementary Figure 1 Cancer Cell Fraction heatmaps obtained from WES data in metastatic CN-low endometrial carcinomas. Heatmaps represent the presence (blue) or absence (grey) of the mutations identified by WES in metastatic endometrial carcinomas molecularly classified as CN-low (EEC1-4). Different blue tones indicate the cancer cell fraction (CCF) value obtained for each variant (see *3.5.4.3*). Loss of heterozigosity (LOH) is represented by a diagonal white line in the corresponding variant box. The boxes of those variants identified as clonal are labeled with a yellow square. EEC: endometrioid endometrial carcinoma, T: primary tumor, M: metastasis.



Supplementary Figure 2 Cancer Cell Fraction heatmaps obtained from WES data in metastatic MSI and serous-like endometrial carcinomas. Heatmaps represent the presence (blue) or absence (grey) of the mutations identified by WES in metastatic endometrial carcinomas molecularly classified as MSI (EEC5-7) or serous-like (SEC1-3). Different blue tones indicate the cancer cell fraction (CCF) value obtained for each variant (see *3.5.4.3*). Loss of heterozigosity (LOH) is represented by a diagonal white line in the corresponding variant box. The boxes of those variants identified as clonal are labeled with a yellow square. EEC: endometrioid endometrial carcinoma, SEC: serous endometrial carcinoma, T: primary tumor, M: metastasis.